Supplementary Materials for

Civic Honesty Around the Globe

Alain Cohn, Michel André Maréchal, David Tannenbaum, Christian Lukas Zünd
correspondence to: adcohn@umich.edu, michel.marechal@econ.uzh.ch

This PDF file includes:

Contents

List of Figures

List of Tables

Materials and Methods

Lost Wallet Experiments

We visited 355 cities in 40 countries and turned in a total of 17,303 wallets between July 2013 and December 2016. Table S1 provides an overview of the study design, including the countries and cities covered, the amount of money included in the wallets, the names on the business cards, the items on the shopping list, and the number of observations. Our study was approved by the Human Subjects Committee of the Faculty of Economics, Business Administration, and Information Technology at the University of Zurich.

Selection of Countries and Cities We selected our sample of countries based on several factors, most important being that the country have a sufficient number of large cities. As a rough guide during the planning process we aimed for populations of at least 100,000, but used this rule flexibly as availability of feasible drop-off locations varied substantially even when restricting ourselves to large cities. In addition to city size, a country had to be relatively easy to visit and safe enough for our research assistants to perform the wallet drop-offs. Customs, immigration, and banking regulation also played a role because research assistants needed to either import or withdraw sufficient money to place in the wallets.

For each country we typically chose five to eight cities to perform the wallet drop-offs. We took the largest cities of a country as a starting point and adapted the list to accommodate safety concerns, cover the main regions of a country, and avoided cities that belong to the same metropolitan area. As cities differed in their size, the number of drop-offs in a city was determined by the relative population size using the following formula:

$$N_i = \frac{\sqrt{POP_i}}{\sum\limits_{k \in C} \sqrt{POP_k}} * N_C^{target}, \tag{1}$$

where N_i is the number of drop-offs in city i, POP_i is city i's population size, k is a city sampled from country C, and N_C^{target} is the target sample size for a given country. This adjustment was designed to avoid a single city dominating our estimates of a country's response rate, while also giving greater weight to more populated cities as they represent a greater fraction of a country's total population (and so tend to be more influential politically, culturally, and economically).

Number of Drop-offs We usually collected 400 observations per country, but there were exceptions to this rule. For some countries we set a different target sample size, and for other countries we ended up with deviations from the targeted sample size due to unforeseen circumstances or minor errors in the data collection process.

We collected a greater number of observations in the US, UK, and Poland since we ran two additional treatment conditions (BigMoney and Money-NoKey conditions) in these countries.[1] In the United States, we collected 300 wallets each in the NoMoney and Money conditions and 200 wallets each in the Money-NoKey and BigMoney conditions, yielding a total sample of 1,000 observations. In the UK, we turned in 200 wallets each in the NoMoney, BigMoney, and Money-NoKey condition, and 600 wallets in the Money condition. We were unable to track email responses for 67 wallets in the Money condition and one wallet in the BigMoney condition due to a procedural error, leaving us with a total of 1,132 observations in the UK. In Poland, we turned in 200 wallets in each of our four conditions, yielding a total sample of 800 observations.

For eight countries — Croatia, Denmark, Ghana, Israel, Kenya, Norway, Serbia, and Russia — we set a sample size target of 300 drop-offs due to either a limited number of sufficiently large cities or due to safety concerns. For India, we made a last minute change by replacing Chennai with Coimbatore due to severe flooding that took place in February 2015. In Kenya we did not carry out data collection in the last city visited (Malindi) because the research assistant was arrested and interrogated by the military police for suspicious activity. In Chile, four wallets had to be excluded

[1] In the US, UK, Poland, France, Italy, and Spain we also conducted additional treatment arms which were orthogonal to our NoMoney and Money conditions. These additional treatment arms mostly involved changing subtle characteristics about the owner of the wallet. We plan to report these results in a separate paper. For France and the UK we observed no significant effect on the reporting rate across these additional conditions, so we pool the data for those two countries here to increase the precision of our estimates. Excluding this additional data from the analysis has virtually no effect on the results we report below.

from the analysis because of a handling mistake which made it impossible to ascertain the location of where the wallets were turned in.

Additional minor deviations from the target sample size occurred due to rounding errors in the allocation of drop-offs to different cities or because experimenters could not find a suitable replacement for a closed drop-off location in time. Countries with minor deviations are marked by a footnote in Table S1.

Selection of Drop-off Locations We focused on five types of institutions as drop-off locations: (i) banks, (ii) theaters, museums, or other cultural establishments, (iii) post offices, (iv) hotels, and (v) police stations, courts of law, and other public offices. While we aimed at an equal distribution of institutions, this was not always feasible. In particular, post offices were sometimes hard to find near city centers as they are often spread over geographic regions. Our final distribution was 23% for banks, 20% for cultural establishments, 14% for post offices, 22% for hotels, and 21% for public offices.

Drop-off locations were always planned in advance. To find appropriate locations, we used official websites (e.g., for police stations), travel guides (e.g., for hotels and museums), and Google Maps. To reduce travel time, we advised research assistants to select drop-off locations close to a city center and to choose drop-off locations within walking distance of each other. To avoid suspicion, we excluded drop-off locations that were next to or across the street from one another. We also advised research assistants to select locations that were far enough away from a given police station as to reduce the risk that multiple recipients would turn in wallets to the same police station. When available, we used Google Street View to verify that a location still existed and that the location was easily accessible from the street. Prior to performing the drop-offs, research assistants also checked for national and local holidays, opening hours, and specific working culture (e.g., siesta in Spain).

The Wallets Our wallets were transparent business card cases (see Fig S1). We used transparent cases to ensure that recipients could inspect the wallet's contents without having to open it. Each

Fig. S1. Example lost wallet

Example of a wallet used in our field experiments. All wallets belonged to a male software developer with country-specific names (see Table S1 for the complete list of names). We placed the business cards in the wallets so that this information was visible to all participants. The wallet dimensions were 93mm x 59mm x 5mm and it weighed approximately 24 grams in the NoMoney condition.

wallet contained the same personal items: (i) three identical business cards, (ii) a grocery list, and (iii) a key. The business cards displayed the owner's name, email address, and job title. Their purpose was to identify the owner and provide contact details.

The business cards and shopping list serve to identify the owner as a local resident, signaling that it would be relatively easy to contact the owner and return the wallet. For the business cards, we typically created three fictitious male owners for each country using common local names. We used several sources to assemble lists of common first and last names, which we then checked to avoid names used as references for generic or unidentified persons (e.g., John Doe), were shared with celebrities, or led to a single user-profile on Facebook. The business cards provided the owner's email address, and identified him as a freelance software engineer (to avoid attempts by recipients to reach the owner through his place of employment).

There were some exceptions to how we generated business cards and shopping lists for our wallets. In Switzerland and the Czech Republic, we used the real name of research assistants so that we would be able to collect reported wallets for our internal validation check. For these two countries we also decided to use only two identities (rather than three) so that we could pick up a larger share of the wallets. In Canada and India, different names were used for some cities to accommodate for the local language. Due to South Africa's history of race relations, we used two discernibly white and two discernibly black names, leading to a total of four names.[2] We made occasional changes to the shopping lists to accommodate local customs, such as using rice instead of pasta or substituting milk with some other beverage where lactose intolerance was common. Table S1 provides a comprehensive list of names and shopping lists.

Drop-off Procedure We recruited eleven male and two female research assistants to perform the drop-offs. All research assistants were recruited from two German speaking universities and born between 1985 and 1993.[3] Research assistants were carefully trained and received detailed manuals on how to carry out the drop-offs. After walking into a building, research assistants were instructed to approach an employee at the counter and say: "*Hi, I found this [showing the wallet] on the street just around the corner.*" Then, they put the wallet on the counter and pushed it over to the employee: "*Somebody must have lost it. I'm in a hurry and have to go. Can you please take care of it?*"[4] The research assistant subsequently left the building without leaving their contact details or asking for a receipt.[5] This interaction was designed to minimize recipients' concerns about being punished, since there was no written proof that a wallet had been turned in. Furthermore, by telling recipients that the wallet was found outside the building around the corner, we avoided possible concerns that the owner might come back and look for the wallet in that exact location.

Experimental Conditions Our primary experimental manipulation varied the amount of money in the wallet. In the "NoMoney" condition, the wallets only contained business cards, a shopping

[2] In South Africa reporting rates were remarkably similar between Black and White names. Reporting rates were always between 32% and 35%, with no significant difference in reporting rates between the four identities ($\chi^2_3 = 0.255$, $P = 0.968$).

[3] In the Robustness Checks section on page 77 we assess the influence of research assistants and find no evidence that differences between experimenters are driving our main results.

[4] Recipients were always approached in English, but research assistants also used a translator app on their cell phones in case a recipient was not conversant in English.

[5] Recipients rarely refused to take the wallet. The median rejection rate was less than 0.4%, with only five countries exhibiting rejection rates above 1% (and none greater than 5%). Columns (1) and (2) in Table S6 shows that rejection rates did not significantly differ between the Money, NoMoney, and Money-NoKey conditions. We find a marginally significant difference ($t_{2884} = 1.77$, $P = 0.077$) between the Money and the BigMoney condition, as shown in column 2. Using χ^2-tests, we find that only 3.3% of all possible pairwise country comparisons are significant at the 5% level after controlling for the false discovery rate (*26*).

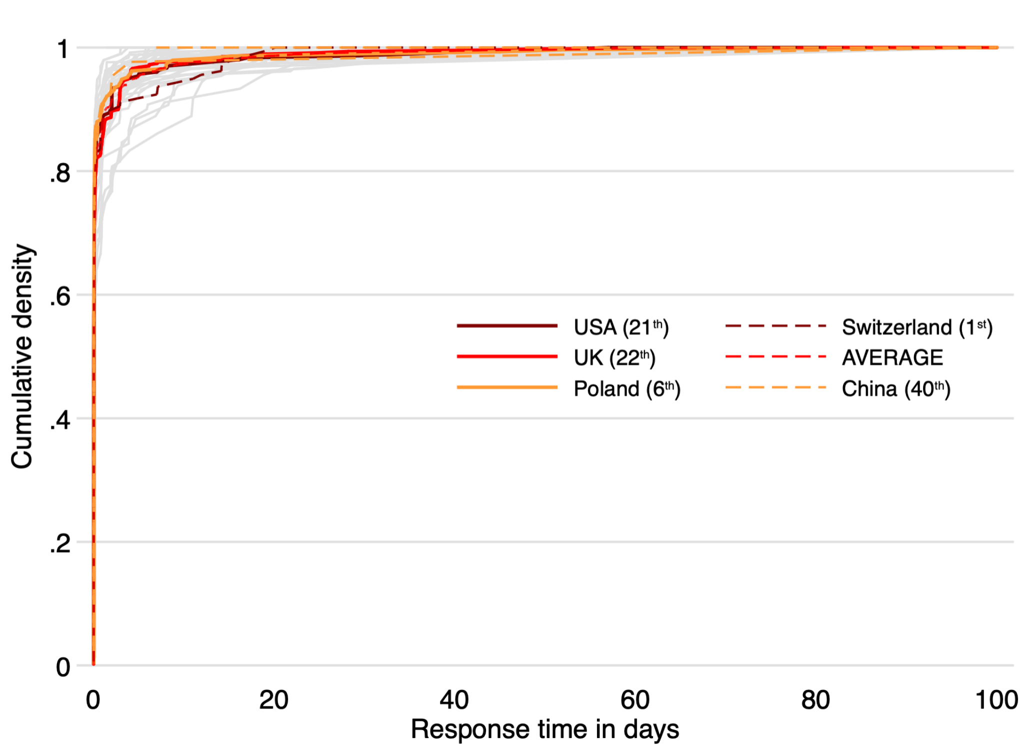
list, and a key. In the “Money” condition, the wallets also contained the equivalent of US $13.45. We used local currencies, and to ensure comparability across countries we adjusted the amounts for purchasing power parity using data from the International Monetary Fund. Table S1 provides the exact amounts of money used in each country.

In three countries (the United Kingdom, Poland, and the United States), we conducted two additional treatment conditions. We ran a “BigMoney” condition that was identical to the Money condition but with the equivalent of US $94.15 in the wallets (i.e., seven times the amount found in the Money condition). We also ran a “Money-NoKey” condition identical to the Money condition but the wallets did not contain a key. Because the key is only valuable to the owner, the Money-NoKey condition only varies the harm caused to the owner relative to the Money condition. This treatment therefore allows us to isolate the role of altruism in people’s decision to return a lost wallet.

We randomly assigned treatments and owner identities to drop-off locations. Tables S2-S5 provide descriptive statistics and demonstrate that individual characteristics and situational factors are well balanced across treatments.

Measuring Civic Honesty Our key outcome measure was whether a recipient contacted the owner to return the wallet. We created our own email server to collect responses. The business cards in each wallet had a unique email address that allowed us to automatically assign incoming emails to its respective drop-off location and to automatically send a reply message in the local language. The following reply message was sent three hours and fifteen minutes after receiving an email from the recipients: “*Hello, thank you very much for your email. I really appreciate your help. Unfortunately, I have already left town. The content of the business card holder and the key are not important to me. You can keep all of it or donate it to charity. Best regards, [firstname] [lastname].*” If present, we specifically mentioned the key because recipients would frequently inquire about the key in follow-up emails. If multiple emails were sent to the same email address then we flagged them for review by a research assistant. The majority of these emails did not

Fig. S2. Cumulative distribution function of response times



Notes: Cumulative distribution function for the time elapsed between the drop-off of the wallets and the email responses from the recipients by country. The three main countries, and the countries with the highest and lowest response rate in the NoMoney condition are highlighted (ranking in parentheses).

necessitate further action.

Besides automating part of the data collection, the use of a private email server allowed us to register attempts to return a wallet even if the email address was spelled incorrectly. As long as the domain name was spelled correctly, a research assistant could manually reassign the email to the correct drop-off location. Common errors involve forgetting the dot in the email address or using a similar name, such as"lars-andersen" or "lars.andresen" instead of "lars.andersen."

We recorded emails that were sent within 100 days after the drop-offs. The median response time was roughly 26 minutes across all countries, and about 88% of emails arrived within 24 hours (see Figure S2). Table S7 shows that response times did not significantly differ across treatments. Moreover, we find little variation in response times between countries. Using two-sample t-tests, we find that only 1.5% of all possible pairwise country comparisons are significant at the 5%

level (FDR-adjusted P-values). Average response times and response rates by country were not significantly correlated (Spearman's $\rho = 0.162$, $P = 0.319$).

Measuring Recipient Characteristics and Situational Factors Upon leaving the locations, our research assistants filled out a short survey to collect additional information about the drop-offs. This data allowed to account for incidental factors that varied across locations. Research assistants recorded the following information:

- *Recipient gender*. Research assistants took note of the recipient's gender, which was coded as 0 for female and 1 for male.

- *Recipient age*. Research assistants estimated the recipient's approximate age on a 6-point scale: < 20, 20-30, 30-40, 40-50, 50-60, > 60. For all analyses using age we use a median split dummy variable in which we coded as 1 if the recipient was estimated to be 40 years or older and 0 otherwise. We used a median split for purposes of simplicity; using a set of indicator variables for each age category does not meaningfully affect any of the treatment effects we report.

- *Busyness*. Research assistants estimated how busy the recipient was on a 7-point scale from "not at all" (0) to "very busy" (6).

- *Local recipient*. Research assistants assessed whether the recipient was a foreigner on a 7-point scale from "local" (0) to "unclear" (3) to "foreigner" (6). We coded this variable as 1 if the recipient was rated below the midpoint of the scale and 0 otherwise. We used this indicator variable for purposes of simplicity; treating this local residency as a continuous variable does not meaningfully affect any of the treatment effects we report.

- *No English*. Whether the research assistant had to use a different language than English to communicate with the recipient (using a mobile phone app). We coded this as 0 if the recipient understood English, and 1 if the recipient did not. This variable was always coded as 0 if a region is English speaking.

- *Recipient understood situation*. Research assistants assessed the extent the recipient understood the situation on a 7-point scale from "not at all" (0) to "fully understood" (6). We only collected this information after finishing data collection in Poland and the UK.

- *Friendliness*. Research assistants assessed the friendliness of the recipient on a 7-point scale from "very unfriendly" (0) to "very friendly" (6). We only collected this information after finishing data collection in Croatia, Greece, Poland, Romania, Serbia, and the UK.

- *Computer*. Research assistants noted if there was a computer at the recipient's desk (0 = computer absent, 1 = computer present).

- *Coworkers*. Research assistants took note of how many employees participated in or closely witnessed the exchange. They had the following response options: one, two, three, or more than three. For all analyses using this variable we coded this variable as 1 if multiple coworkers participated or closely witnessed the exchange and 0 otherwise. We used this indicator variable for purposes of simplicity; using a set of indicator variables for each response option does not meaningfully affect any of the treatment effects we report.

- *Other bystanders*. Research assistants took note of how many other people could witness the drop-off. They had the following response options: none, fewer than five, five or more. For all analyses using presence of bystanders we coded this variable as 1 if any bystanders were present and 0 otherwise. We used this indicator variable for purposes of simplicity; using a set of indicator variables for each response option does not meaningfully affect any of the treatment effects we report.

- *Security camera*. Research assistants took note of whether a security camera was visible in the room (0 = no camera visible, 1 = camera visible). We only collected this information after finishing data collection in Poland and the UK.

- *Security guard*. Research assistants took note of whether a security was present (0 = guard present, 1 = guard absent). We only collected this information after finishing data collection in Croatia, Greece, Poland, Romania, Serbia, and the UK.

Country-level Correlates of Civic Honesty As a supplement to our experimental study, we also examined country-level predictors of civic honesty. We examined how rates of civic honesty vary according to the following set of country-level characteristics:

- *Country GDP*. Logarithm of country gross domestic product based on purchasing-power-parity per capita in 2010 from the IMF World Economic Outlook (*27*).
- *Log. soil fertility*. Logarithm of soil suitability, obtained from (*28*). The data is originally from Ramankutty *et al.* (*29*) who estimated soil suitability at half-degree resolution based on soil pH and soil carbon density. The data was then aggregated at the country-level by (*30*). Missing data on Serbia has been replaced with data from Yugoslavia.
- *Log. abs. latitude*. Logarithm of the absolute latitude of a country's approximate geodesic centroid, obtained from (*28*). The data is originally from the CIA's *World Factbook*. Missing data on Serbia has been replaced with data from Yugoslavia.
- *Distance to waterway*. Distance (in 100 km) to the nearest ice-free coastline or sea-navigable river, obtained from (*28*). The data is originally from (*31*). Missing data on Serbia has been replaced with data from Yugoslavia.
- *Temperature*. Average monthly temperature (in Celsius degrees) of a country between 1961 and 1990, obtained from (*28*). The data is originally from the G-ECON project (*32*). Missing data on Serbia has been replaced with data from Yugoslavia.
- *Precipitation*. Average monthly precipitation (in mm per month) of a country between 1961 and 1990, obtained from (*28*). The data is originally from the G-ECON project (*32*). Missing data on Serbia has been replaced with data from Yugoslavia.
- *Mean elevation*. Mean elevation of a country (in km) above sea level, obtained from (*28*). The data is originally from the G-ECON project (*32*). Missing data on Serbia has been

replaced with data from Yugoslavia.

- *Terrain roughness*. Degree of terrain roughness, obtained from (*28*). The data is originally from the G-ECON project (*32*). Missing data on Serbia has been replaced with data from Yugoslavia.

- *Temperature (Volatility)*. Ancestry adjusted volatility of temperature between 1900 and 2000. Based on the *Climatic Research Unit (CRU)* database, and constructed using the method outlined in (*33*). The variable is obtained from (*34*). Missing data on Serbia has been replaced with data from Yugoslavia.

- *Precipitation (Volatility)*. Ancestry adjusted volatility of precipitation between 1900 and 2000. Based on the *Climatic Research Unit (CRU)* database, and constructed using the method outlined in (*33*). The variable is obtained from (*34*). Missing data on Serbia has been replaced with data from Yugoslavia.

- *Pathogen prevalence*. Historic prevalence of nine infectious diseases (leishmanias, schistosomes, trypanosomes, leprosy, malaria, typhus, filariae, dengue, and tuberculosis) based on epidemiological atlases from the first half of the 20th century, as constructed by (*35*). The variable is obtained from (*36*). Data on Kazakhstan (and several other countries not covered by our study) has been fitted based on an index of seven pathogens (excluding leprosy and tuberculosis) (*36*).

- *Pronoun drop not allowed*. Share of individuals that speak a language that does not allow dropping the first-person pronoun (i.e., "I"), thereby putting more emphasis on the individual (*37*). The variable was obtained from (*37*). The data is originally from (*38*). Data on Croatia, Kazakhstan, Morocco, and Serbia has been manually completed based on the major languages using data from the *World Atlas of Language Structures (WALS)*.

- *Politeness distinction*. Share of individuals that speak a language that prescribes the use of different pronouns (e.g., “tu" and “vous" in French) depending on the relationship between the speakers. This is a trait that has been linked to hierarchy and power distance (*37*). The variable was obtained from (*37*). The data is originally from (*38*). Data on Croatia, Kazakhstan, Morocco, and Serbia has been manually completed based on the major languages using data from the *World Atlas of Language Structures (WALS)*.

- *Weak future time reference*. Share of individuals that speak a language with a weak future time reference, obtained from (*39*). Languages with a weak future time reference allow the speaker to use the same grammatical tense to speak about present and future events instead of having a grammatically distinct future tense. Data on Brazil, Morocco, Peru, Serbia, South Africa, Indonesia, Ghana, Kenya, Kazakhstan, India, and the United Arab Emirates have been manually completed based on the major languages using data from the *World Atlas of Language Structures (WALS)*.

- *Share of protestants*. Percentage of a country's population that is protestant, obtained from (*28*). The data is originally from (*40*). Missing data on Serbia has been manually completed using data from the Serbian census in 2002.

- *Family ties*. Strength of family ties calculated following Alesina and Giuliano (*41*). The variable is the first principal component of three family-related questions in the *World Value Survey (WVS)*: (i) "For each of the following, indicate how important it is in your life. - Family:", on a 4-point scale from 1 (*not important at all*) to 4 (*very important*), (ii) "With which of these two statements do you tend to agree? A: One does not have the duty to respect and love parents who have not earned it; B: Regardless of what the qualities and faults of one's parents are, one must always love and respect them." (iii) "Which of the following statements best describes your views about parents' responsibilities to their children? A: Parents have a life of their own and should not be asked to sacrifice their own well-being for the sake of their children; B: It is the parents' duty to do their best for their children even at the expense of their own well being."

- *State history*. State history index (*42*). For each period of 50 years from year 1 C.E. to 1950, a country's experience with supra-tribal government is coded for (i) the existence of a government above the tribal level, (ii) whether said government was foreign or locally based, and (iii) how much of the current country it ruled. A discount factor of 5% for each 50 years is applied so that more recent experience with statehood are weighted more heavily in the index. The variable is obtained from (*43*).

- *Years of democracy*. Years since the polity score in the *Polity IV* data set is strictly above zero, starting from 1800 or the year of independence for countries that became independent later. The polity score is defined by subtracting the autocracy score from the democracy score and ranges from "strongly democratic" (10) to "strongly autocratic" (–10).

- *Executive constraints*. Constraints on executive scale from the *Polity IV* data set. The scale takes values from "unlimited authority" (1) to "executive parity or subordination" (7), the later being defined as a situation in which "accountability groups have effective authority equal to or greater than the executive in most areas of activity."

- *Judicial independence*. Judicial independence as of 1995, obtained from (*44*). The data is originally from LaPorta *et al.* (*45*) who defined the variable as the sum of three sub-scales measuring (i) tenure of supreme court judges, (ii) tenure of the highest ranked judges ruling on administrative cases, and (iii) the existence of case law. The variable is normalized to range from zero to one.

- *Constitutional review*. Constitutional review as of 1995, obtained from (*44*). The data is originally from LaPorta *et al.* (*45*) who defined the variable as the sum of two sub-scales measuring (i) the extent to which judges of the supreme or constitutional court can review the constitutionality of laws and (ii) how difficult it is to change the constitution. The variable is normalized to range from zero to one.

- *Electoral rule: Plurality*. Percentage of years between 1975 and 2000 in which a first-past-the-post or winner-takes-all system was used to elect legislators, obtained from (*44*). The data is originally from (*46*).

- *Electoral rule: Proportionality*. Percentage of years between 1975 and 2000 in which a proportional system was used to elect legislators, i.e., legislators were elected based on the share of votes that their party received in an election. The variables is obtained from (*44*). The data is originally from (*46*).

- *Primary education 1920*. Primary school enrollment in 1920, obtained from (*47*).

Tab. S1. Sample overview

Country	Cities	Treatments	Names	Shopping List	Languages Email	N
Argentina (16 Jul. 2015 – 8 Aug. 2015)	Buenos Aires Córdoba Mar del Plata Mendoza Rosario Salta San Miguel de Tucumán Santa Fe	Money (ARS 48.50) NoMoney	Eduardo Martinez Guillermo Garcia Ignacio Lombardi	dulce de leche pan costilla lata duraznos	Spanish	400
Australia (4 May 2015 – 25 May 2015)	Adelaide Brisbane Canberra Goldcoast Melbourne Newcastle Perth Sydney	Money (AUD 20) NoMoney	Jack Williams James Smith William Jones	milk bread noodles bananas	English	399[a]
Brazil (25 Jul. 2015 – 13 Aug. 2015)	Belo Horizonte Brasília Curitiba Fortaleza Manaus Rio de Janeiro Salvador São Paulo	Money (BRL 21.75) NoMoney	Gabriel Pereira Lima Lucas de Oliveira Souza Rafael da Luz Santos	leite pão macarrão frutas	Portuguese	399[a]
Canada (10 Sep. 2015 – 6 Oct. 2015)	Calgary Edmonton Ottawa Toronto Vancouver	Money (CAD 16.50) NoMoney	David Smith Jacob Brown Robert Wilson	milk bread pasta bananas	English	400
	Montréal Québec		Alexandre Gagnon Olivier Roy Thomas Tremblay	lait pain pâtes bananes	French	

Continued

Tab. S1. Sample Overview (continued)

Country	Cities	Treatments	Names	Shopping List	Languages Email	N
Chile (16 Nov. 2016 – 13 Dec. 2016)	Antofagasta Concepción La Serena Rancagua Santiago de Chile Talca Temuco Valparaíso	Money (CLP 4650) NoMoney	Diego Rojas Francisco Muñoz Jorge Gonzalez	leche pan choclo plátanos	Spanish	396^{b}
China (7 Jul. 2015 – 27 Jul. 2015)	Beijing Chengdu Guangzhou Hangzhou Shanghai Shenzhen Tianjin Xi'an	Money (RMB 49) NoMoney	Chang Wei Li Qiang Wang Lei	矿泉水 包子 方便面 苹果	Chinese	400
Croatia (25 Mar. 2014 – 4 Apr. 2014)	Osijek Rijeka Split Zadar Zagreb	Money (HRK 53.50) NoMoney	Ivan Kovačević Marko Horvat Tomislav Babić	mlijeka kruha tjestenine banane	Croatian	300
Czech Republic (21 Oct. 2014 – 7 Nov. 2014)	Brno Liberec Olomóc Ostrava Plzeň Praha	Money (CZK 170) NoMoney	Marek Pospíšil Václav Korbel $-^{c}$	mléka chleba těstovin banány	Czech	400
Denmark (15 Jul. 2014 – 25 Jul. 2014)	Aalborg Århus København Odense	Money (DKK 109) NoMoney	Christian Rasmussen Lars Andersen Søren Jensen	mælk brød pasta bananen	Danish	300

Continued

Tab. S1. Sample Overview (continued)

Country	Cities	Treatments	Names	Shopping List	Languages Email	N
France (21 Jul. 2014 – 26 Aug. 2014)	Angers Bordeaux Brest Le Havre Lille Lyon Marseille Montpellier Nantes Nice Paris Rennes Saint-Étienne Strasbourg Toulon Toulouse	Money (EUR 11.50) NoMoney	Benoît Barbier Bertrand Moreau Philippe Richard	lait pain pâtes bananes	French	802[a]
Germany (23 Jun. 2014 – 11 Jul. 2014)	Berlin Köln Dresden Frankfurt Hamburg Leipzig München Stuttgart	Money (EUR 10.50) NoMoney	Michael Weber Stefan Richter Thomas Becker	Milch Brot Pasta Bananen	German	400
Ghana (30 Aug. 2016 – 15 Sep. 2016)	Kumasi Sekondi Sunyani Accra Ashaiman	Money (GHA 12.25) NoMoney	Daniel Mensah Emmanuel Ansah Samuel Owusu	Nsuo Aburo ℰmo Kwadu	English	299[a]
	Tamale			bihim kahu siŋkafa kodu		

Continued

Tab. S1. Sample Overview (continued)

Country	Cities	Treatments	Names	Shopping List	Languages Email	N
Greece (25 Apr. 2014 – 14 May 2014)	Athens Chania Heraklion Patras Thessaloniki	Money (EUR 9.00) NoMoney	George Nikolaidis Giannis Papadopoulos Kostas Vlachos	ψωμί ντομάτες λάδι ζυμαρικά	Greek	400
India (22 Nov. 2015 – 10 Dec. 2015)	New Delhi Ahmedabad Jaipur Mumbai	Money (INR 230) NoMoney	Gaurav Kapoor Govind Malik Shekhar Kohli	milk bread rice bananas	English	400
	Hyderabad Bangalore Coimbatore[d]		Kapil Prasad Kesav Kumar Niraj Das			
	Kolkata		Pratyush Chatterjee Pravin Patil Vinod Ghosh			
Indonesia (8 Apr. 2016 – 4 May 2016)	Bandung Batam Jakarta Makassar Medan Palembang Semarang Surabaya	Money (IDR 51300) NoMoney	Budi Nugraha Surya Adi Tono Hendrianta	minyak goring bj telur beras pisang	Indonesian	400
Israel (13 Jul. 2015 – 2 Aug. 2015)	Ashdod Beersheba Haifa Jerusalem Netanya Tel Aviv	Money (ILS 55) NoMoney	David Mizrahi Itai Cohen Joseph Levy	חלב לחם שחור ביצים תפוחים	Hebrew	300

Continued

Tab. S1. Sample Overview (continued)

Country	Cities	Treatments	Names	Shopping List	Languages Email	N
Italy (15 Jun. 2014 – 11 Jul. 2014)	Bari Bologna Catania Firenze Genova Messina Milano Napoli Padova Palermo Roma Taranto Torino Trieste Venezia Verona	Money (EUR 10.75) NoMoney	Antonio Gallo Giuseppe Russo Roberto Bianchi	latte pane pasta banane	Italian	400
Kazakhstan (4 Jun. 2016 – 16 Jun. 2016)	Almaty Astana Karaganda Oskemen Pavlodar Semey Shymkent Taraz	Money (KZT 1203) NoMoney	Alexey Omarov Bekzat Akhmetov Kirill Ospanov	сүт нан кеспе банан	Kazakh	400
Kenya (10 Nov. 2015 – 24 Nov. 2015)	Nairobi Nakuru Nyeri	Money (KES 507) NoMoney	John Omondi Peter Kihiga Samuel Jabali	iria ngima chapati mũraru	English	274[e]
	Eldoret Kisumu			chak kuon chapati rabolo		
	Mombasa –[e]			maziwa ugali chapati ndizi		

Continued

Tab. S1. Sample Overview (continued)

Country	Cities	Treatments	Names	Shopping List	Languages Email	N
Malaysia (23 May 2016 – 13 Jun. 2016)	George Town Ipoh Johor Bahru Kota Kinabalu Kuala Lumpur Melaka Petaling Jaya Shah Alam	Money (MYR 10.15) NoMoney	Amir bin Roslan Farid bin Azhari Ramlee bin Khadir	susu telor beras asam keping	Malay	400
Mexico (7 Sep. 2015 – 28 Sep. 2015)	Chihuahua Guadalajara León Mérida Mexico City Monterrey Puebla Tijuana	Money (MXN 105) NoMoney	Carlos García Daniel Martínez José Hernandez	leche pan pasta plátanos	Spanish	400
Morocco (25 May 2015 – 12 Jun. 2015)	Agadir Casablanca Fez Queneitra Marrakesh Meknès Rabat Tangier Tétouan	Money (MAD 49) NoMoney	Ahmed el Mernissi Mohamed Bennani Yassine ben Aissa	حليب خبز الكسكس موز	French	402[a]
Netherlands (12 Aug. 2014 – 26 Aug. 2014)	Amsterdam Eindhoven Groningen Rotterdam Den Haag Tilburg Utrecht	Money (EUR 10.80) NoMoney	Jan de Vries Martijn de Jong Sander Jansen	melk brood pasta bananen	Dutch	400

Continued

Tab. S1. Sample Overview (continued)

Country	Cities	Treatments	Names	Shopping List	Languages Email	N
New Zealand (7 Apr. 2015 – 28 Apr. 2015)	Auckland Christchurch Dunedin Hamilton Napier Tauranga Wellington	Money (NZD 20) NoMoney	Jack Taylor Josua Brown Liam Williams	milk bread noodles bananas	English	400
Norway (27 Apr. 2015 – 12 May 2015)	Bergen Kristiansand Oslo Stavanger Trondheim	Money (NOK 125) NoMoney	Bjorn Hansen Jan Johansen Per Olsen	melk brød nudler bananer	Bokmål	300
Peru (1 Aug. 2016 – 19 Aug. 2016)	Arequipa Chiclayo Cuzco Iquitos Lima Piura Trujillo	Money (20.50) NoMoney	Carlos Sánchez Días José García González Luis Flores Ramírez	leche pan arroz plátanos	Spanish	400

Continued

Tab. S1. Sample Overview (continued)

Country	Cities	Treatments	Names	Shopping List	Languages Email	N
Poland (25 Jul. 2013 – 29 Aug. 2013)	Białystok Bydgoszcz Bytom Częstochowa Gdańsk Gdynia Gliwice Katowice Kielce Krakow Łódć Lublin Opole Poznan Radom Sosnowiec Szczecin Toruń Warszawa Wrocław	BigMoney (PLN 175) Money (PLN 25) Money-NoKey (PLN 25) NoMoney	Edward Kowalski Marek Nowak Paweł Wiśnieswski	mleka chleb makaron banany	Polish	800
Portugal (15 May 2014 – 31 May 2014)	Braga Coimbra Faro Lisboa Porto Setúbal	Money (EUR 8.50) NoMoney	João Silva Fernandes Miguel Ferreira Rodrigues Rodrigo Santos Pereira	leite pão pasta bananas	Portuguese	400
Romania (25 Mar. 2014 – 15 Apr. 2014)	Brașov București Cluj-Napoca Constanța Galați Iași Timișoara	Money (RON 28) NoMoney	Andrei Popescu Constantin Radu Gheorghe Matei	lapte pâine paste banane	Romanian	400

Continued

Tab. S1. Sample Overview (continued)

Country	Cities	Treatments	Names	Shopping List	Languages Email	N
Russia (4 Aug. 2015 – 18 Aug. 2015)	Kazan Moscow Nizhny Novgorod Novosibirsk Omsk St.Petersburg Yekaterinburg	Money (RUB 258) NoMoney	Daniil Smirnov Dimitri Ivanov Ivan Kuznetsov	молоко хлеб лапша бананы	Russian	302[a]
Serbia (7 Apr. 2014 – 17 Apr. 2014)	Belgrade Kragujevac Niš Novi Sad Subotica	Money (DIN 612) NoMoney	Dragan Pavlović Nikola Stojanović Vladimir Nikolic	млека хлеб тестенина банане	Serbian	300
South Africa (13 Jan. 2016 – 11 Feb. 2016)	Bloemfontein Cape Town Durban East London Johannesburg Pietermaritzburg Port Elizabeth Pretoria	Money (ZAR 69) NoMoney	Johan Fourie Michael Botha Thabo Molefe Tshepo Mokwena	milk bread rice bananas	English	399[a]
Spain (13 May 2014 – 25 Jun. 2014)	A Coruña Alicante Barcelona Bilbao Córdoba Gijón Madrid Málaga Murcia Palma Sevilla Valencia Valladolid Vigo Zaragoza	Money (EUR 9.50) NoMoney	Antonio García González José Fernández García Manuel González Fernández	leche pan pasta plátanos	Spanish	400

Continued

Tab. S1. Sample Overview (continued)

Country	Cities	Treatments	Names	Shopping List	Languages Email	N
Sweden (21 Aug. 2014 – 5 Sep. 2014)	Göteborg Helsingborg Jönköping Linköping Lund Malmö Norrköping Stockholm Uppsala	Money (SEK 115) NoMoney	Anders Johansson Lars Andersson Mikael Karlsson	mjölk bröd pasta bananer	Swedish	400
Switzerland (8 Sep. 2014 – 26 Sep. 2014)	Basel Bern Luzern St. Gallen Winterthur Zürich	Money (CHF 20.75) NoMoney	Daniel Martin Marco Schwarz – [c]	Milch Brot Pasta Bananen	German	399[a]
	Geneva Lausanne			lait pain pâtes bananes	French	
Thailand (16 May 2015 – 12 Jun. 2015)	Bangkok Chiang Mai Hat Yai Khon Kaen Nakhon Ratchasima Nakhon Si Thammarat Ubon Ratchathani Udon Thani	Money (THB 166) NoMoney	Charoen Bongkot Somsak Banyat Thongchai Malechan	พริกเม็ดใหญ่ ซอสถั่วเหลือง ข้าว มะม่วง	Thai	400
Turkey (26 Jun. 2014 – 11 Jul. 2014)	Adana Ankara Antalya Gaziantep İstanbul İzmir Konya	Money (TRY 16) NoMoney	Hakan Kaya Mesut Demir Mustafa Şahin	süt ekmek makarna muz	Turkish	400

Continued

Tab. S1. Sample Overview (continued)

Country	Cities	Treatments	Names	Shopping List	Languages Email	N
UAE (4 Nov. 2015 – 26 Nov. 2015)	Abu Dhabi Ajman Al Ain Dubai Fujairah Ras al-Khaimah Sharjah	Money (AED 35.75) NoMoney	Ali al Shehhi Mayed al Shamsi Mohammed al Hammadi	قارورة ماء موز سنيكرز علكة	Arabic	400
UK (4 Sep. 2013 – 7 Oct. 2013)	Aberdeen Belfast Birmingham Brighton Bristol Cardiff Derby Edinburgh Glasgow Hull Leeds Leicester Liverpool London Manchester Newcastle upon Tyne Nottingham Oxford Portsmouth Reading Sheffield Southampton York	BigMoney (GBP 58.50) Money (GBP 8.50) Money-NoKey (GBP 8.50) NoMoney	David Brown Mark Smith Michael Wilson	milk bread pasta bananas	English	1132[f]

Continued

Tab. S1. Sample Overview (continued)

Country	Cities	Treatments	Names	Shopping List	Languages Email	N
US (17 Aug. 2015 – 17 Oct. 2015)	Albuquerque Boston Charlotte Chicago Columbus Denver Houston Indianapolis Jacksonville Las Vegas Los Angeles Louisville Memphis Milwaukee Nashville New York Oklahoma City Philadelphia Phoenix Portland San Antonio San Diego Seattle Tucson Washington, D.C	BigMoney (USD 94.15) Money (USD 13.45) Money-NoKey (USD 13.45) NoMoney	Brad O'Brien Brett Miller Connor Baker	milk bread pasta bananas	English	1000
40 Countries	355 Cities					17303

Notes:
[a] There are minor deviations (+/– 2) from the target sample size in some countries due to rounding errors in the planning of the sample size for each city or because no suitable replacement for a closed drop-off location could be found in time.
[b] A handling mistake made it impossible to identify the drop-off locations of four wallets.
[c] We used only two instead of three names in the Czech Republic and Switzerland. We used real names in these countries to be able to pick up the wallets as part of the internal validation.
[d] Coimbatore was chosen as a last minute replacement for Chennai after the February 2015 flooding.
[e] We exclude the last city that we visited (Malindi) because the experimenter was arrested and interrogated by the military police for suspicious activity.
[f] We excluded several observations because of a printing error concerning the email address on the business cards.

Tab. S2. Descriptive statistics and randomization check for the United Kingdom

	NoMoney		**Money**		**Money-NoKey**		**BigMoney**		**Total sample**		
	mean	*SD*	mean	*SD*	mean	*SD*	mean	*SD*	mean	*SD*	*P*-value
Age 40+	0.335	(0.473)	0.326	(0.469)	0.355	(0.480)	0.382	(0.487)	0.343	(0.475)	0.538
Male	0.385	(0.488)	0.381	(0.486)	0.360	(0.481)	0.357	(0.480)	0.374	(0.484)	0.890
Computer	0.865	(0.343)	0.805	(0.397)	0.810	(0.393)	0.819	(0.386)	0.819	(0.385)	0.298
Coworkers	0.195	(0.397)	0.189	(0.392)	0.195	(0.397)	0.211	(0.409)	0.195	(0.397)	0.934
Other bystanders	0.725	(0.448)	0.775	(0.418)	0.795	(0.405)	0.809	(0.394)	0.776	(0.417)	0.199
Hotel	0.230	(0.422)	0.236	(0.425)	0.250	(0.434)	0.261	(0.440)	0.242	(0.429)	0.868
Bank	0.250	(0.434)	0.248	(0.432)	0.255	(0.437)	0.221	(0.416)	0.245	(0.430)	0.857
Cultural	0.225	(0.419)	0.242	(0.429)	0.225	(0.419)	0.216	(0.413)	0.231	(0.422)	0.875
Public	0.190	(0.393)	0.178	(0.383)	0.175	(0.381)	0.196	(0.398)	0.183	(0.387)	0.928
Postal	0.105	(0.307)	0.096	(0.294)	0.095	(0.294)	0.106	(0.308)	0.099	(0.299)	0.964
Local recipient	0.905	(0.294)	0.927	(0.261)	0.915	(0.280)	0.930	(0.256)	0.921	(0.269)	0.739
No English	0.000	(0.000)	0.000	(0.000)	0.000	(0.000)	0.000	(0.000)	0.000	(0.000)	1.000
Busy	2.145	(1.769)	2.201	(1.900)	2.180	(1.709)	2.166	(1.844)	2.181	(1.832)	0.993
Observations	200		533		200		199		1,132		

Notes: We coded "Age 40+" as 1 if the recipient was judged to be 40 years or older, and 0 otherwise. "Male" was coded as 1 if the recipient was male and 0 otherwise. "Computer" was coded as 1 if there was a computer present at the recipient's desk, and 0 otherwise. "Coworkers" was coded as 1 if a recipient's coworkers participated in the exchange, and 0 otherwise. "Other bystanders" was coded as 1 if any bystanders witness the exchange and 0 otherwise. The variables "Hotel," "Bank," "Cultural," "Public," and "Postal" represent the five types of institutions in which the experiments were performed. "Local recipient" was coded as 1 if the recipient was judged to be a local resident, and 0 otherwise. "No English" was coded as 1 if the recipient was not able to communicate with the research assistant in English, and 0 otherwise. "Busy" was a rating of how busy the recipient was when the wallet was turned in, on a scale from 0 (*not at all*) to 6 (*very busy*). The last column presents *P*-values for the null hypothesis of perfect randomization (χ^2-tests, except for "busy" where we perform a Kruskal-Wallis *H* test).

Tab. S3. Descriptive statistics and randomization check for Poland

	NoMoney		Money		Money-NoKey		BigMoney		Total sample		
	mean	*SD*	mean	*SD*	mean	*SD*	mean	*SD*	mean	*SD*	*P*-value
Age 40+	0.263$^{(2)}$	(0.441)	0.286$^{(1)}$	(0.453)	0.275	(0.448)	0.350$^{(3)}$	(0.478)	0.293$^{(6)}$	(0.456)	0.226
Male	0.288$^{(2)}$	(0.454)	0.352$^{(1)}$	(0.479)	0.265	(0.442)	0.313$^{(2)}$	(0.465)	0.304$^{(5)}$	(0.460)	0.272
Computer	0.894$^{(2)}$	(0.309)	0.839$^{(1)}$	(0.368)	0.905	(0.294)	0.864$^{(2)}$	(0.344)	0.875$^{(5)}$	(0.330)	0.181
Coworkers	0.268$^{(2)}$	(0.444)	0.286$^{(1)}$	(0.453)	0.360	(0.481)	0.333$^{(2)}$	(0.473)	0.312$^{(5)}$	(0.464)	0.173
Other bystanders	0.763$^{(2)}$	(0.427)	0.673$^{(1)}$	(0.470)	0.665	(0.473)	0.732$^{(2)}$	(0.444)	0.708$^{(5)}$	(0.455)	0.095
Hotel	0.175	(0.381)	0.155	(0.363)	0.155	(0.363)	0.120	(0.326)	0.151	(0.359)	0.485
Bank	0.340	(0.475)	0.370	(0.484)	0.375	(0.485)	0.345	(0.477)	0.357	(0.480)	0.848
Cultural	0.165	(0.372)	0.145	(0.353)	0.150	(0.358)	0.175	(0.381)	0.159	(0.366)	0.837
Public	0.190	(0.393)	0.210	(0.408)	0.200	(0.401)	0.230	(0.422)	0.207	(0.406)	0.786
Postal	0.130	(0.337)	0.120	(0.326)	0.120	(0.326)	0.130	(0.337)	0.125	(0.331)	0.980
Local recipient	0.895	(0.307)	0.930	(0.256)	0.950	(0.218)	0.900	(0.301)	0.919	(0.273)	0.144
No English	0.515	(0.501)	0.625	(0.485)	0.560	(0.498)	0.605	(0.490)	0.576	(0.494)	0.116
Busy	2.827$^{(3)}$	(1.964)	2.392$^{(1)}$	(1.906)	2.407$^{(1)}$	(1.912)	2.601$^{(2)}$	(1.961)	2.556$^{(7)}$	(1.940)	0.096
Observations	200		200		200		200		800		

Notes: We coded "Age 40+" as 1 if the recipient was judged to be 40 years or older, and 0 otherwise. "Male" was coded as 1 if the recipient was male and 0 otherwise. "Computer" was coded as 1 if there was a computer present at the recipient's desk, and 0 otherwise. "Coworkers" was coded as 1 if a recipient's coworkers participated in the exchange, and 0 otherwise. "Other bystanders" was coded as 1 if any bystanders witness the exchange and 0 otherwise. The variables "Hotel," "Bank," "Cultural," "Public," and "Postal" represent the five types of institutions in which the experiments were performed. "Local recipient" was coded as 1 if the recipient was judged to be a local resident, and 0 otherwise. "No English" was coded as 1 if the recipient was not able to communicate with the research assistant in English, and 0 otherwise. "Busy" was a rating of how busy the recipient was when the wallet was turned in, on a scale from 0 (*not at all*) to 6 (*very busy*). The last column presents *P*-values for the null hypothesis of perfect randomization (χ^2-tests, except for "busy" where we perform a Kruskal-Wallis *H* test). There are a few missing answers in the drop-off survey. Numbers in superscript parenthesis indicate the number of drop-offs that are missing in each condition.

Tab. S4. Descriptive statistics and randomization check for the United States

	NoMoney		**Money**		**Money-NoKey**		**BigMoney**		**Total sample**		
	mean	*SD*	mean	*SD*	mean	*SD*	mean	*SD*	mean	*SD*	*P*-value
Age 40+	0.487	(0.501)	0.453	(0.499)	0.465	(0.500)	0.450	(0.499)	0.465	(0.499)	0.823
Male	0.377	(0.485)	0.417	(0.494)	0.385	(0.488)	0.535	(0.500)	0.422	(0.494)	0.003
Computer	0.860	(0.348)	0.890	(0.313)	0.910	(0.287)	0.865	(0.343)	0.880	(0.325)	0.314
Coworkers	0.220	(0.415)	0.283	(0.451)	0.290	(0.455)	0.275	(0.448)	0.264	(0.441)	0.223
Other bystanders	0.710	(0.455)	0.720	(0.450)	0.700	(0.459)	0.725	(0.448)	0.714	(0.452)	0.943
Hotel	0.230	(0.422)	0.207	(0.406)	0.270	(0.445)	0.230	(0.422)	0.231	(0.422)	0.438
Bank	0.237	(0.426)	0.223	(0.417)	0.185	(0.389)	0.220	(0.415)	0.219	(0.414)	0.586
Cultural	0.247	(0.432)	0.250	(0.434)	0.220	(0.415)	0.210	(0.408)	0.235	(0.424)	0.671
Public	0.193	(0.396)	0.190	(0.393)	0.270	(0.445)	0.255	(0.437)	0.220	(0.414)	0.067
Postal	0.093	(0.291)	0.130	(0.337)	0.055	(0.229)	0.085	(0.280)	0.095	(0.293)	0.041
Security camera	0.783	(0.413)	0.830	(0.376)	0.835	(0.372)	0.835	(0.372)	0.818	(0.386)	0.322
Security guard	0.173	(0.379)	0.210	(0.408)	0.255	(0.437)	0.220	(0.415)	0.210	(0.408)	0.172
Local recipient	0.947	(0.225)	0.957	(0.204)	0.945	(0.229)	0.945	(0.229)	0.949	(0.220)	0.912
No English	0.000	(0.000)	0.000	(0.000)	0.000	(0.000)	0.000	(0.000)	0.000	(0.000)	1.000
Understood situation	5.953	(0.438)	5.990	(0.173)	6.000	(0.000)	5.970	(0.316)	5.977	(0.294)	0.264
Busy	1.333	(1.661)	1.400	(1.638)	1.145	(1.545)	1.130	(1.447)	1.275	(1.592)	0.230
Observations	300		300		200		200		1,000		

Notes: We coded "Age 40+" as 1 if the recipient was judged to be 40 years or older, and 0 otherwise. "Male" was coded as 1 if the recipient was male and 0 otherwise. "Computer" was coded as 1 if there was a computer present at the recipient's desk, and 0 otherwise. "Coworkers" was coded as 1 if a recipient's coworkers participated in the exchange, and 0 otherwise. "Other bystanders" was coded as 1 if any bystanders witness the exchange and 0 otherwise. The variables "Hotel," "Bank," "Cultural," "Public," and "Postal" represent the five types of institutions in which the experiments were performed. "Security camera" was coded as 1 if a security camera was visible during the exchange, and 0 otherwise. "Security guard" was coded as 1 if a security guard was present, and 0 otherwise. "Local recipient" was coded as 1 if the recipient was judged to be a local resident, and 0 otherwise. "No English" was coded as 1 if the recipient was not able to communicate with the research assistant in English, and 0 otherwise. "Understood situation" was a rating of whether the recipient understood the situation on a scale from 0 (*not at all*) to 6 (*fully understood*). "Busy" was a rating of how busy the recipient was when the wallet was turned in, on a scale from 0 (*not at all*) to 6 (*very busy*). The last column presents *P*-values for the null hypothesis of perfect randomization (χ^2-tests, except for "understood situation" and "busy" where we perform a Kruskal-Wallis *H* test).

Tab. S5. Descriptive statistics and randomization check for the global data

	NoMoney		Money		Total sample		
	mean	*SD*	mean	*SD*	mean	*SD*	*P*-value
Age 40+	$0.408^{(2)}$	(0.491)	$0.411^{(2)}$	(0.492)	$0.409^{(4)}$	(0.492)	0.684
Male	$0.464^{(2)}$	(0.499)	$0.463^{(1)}$	(0.499)	$0.463^{(3)}$	(0.499)	0.967
Computer	$0.773^{(2)}$	(0.419)	$0.775^{(3)}$	(0.418)	$0.774^{(5)}$	(0.418)	0.778
Coworkers	$0.332^{(2)}$	(0.471)	$0.343^{(1)}$	(0.475)	$0.338^{(3)}$	(0.473)	0.155
Other bystanders	$0.645^{(2)}$	(0.479)	$0.656^{(2)}$	(0.475)	$0.651^{(4)}$	(0.477)	0.112
Hotel	0.218	(0.413)	0.217	(0.412)	0.218	(0.413)	0.855
Bank	0.233	(0.423)	0.231	(0.422)	0.232	(0.422)	0.806
Cultural	0.199	(0.400)	0.202	(0.401)	0.201	(0.401)	0.709
Public	0.210	(0.408)	0.211	(0.408)	0.211	(0.408)	0.926
Postal	0.139	(0.346)	0.139	(0.346)	0.139	(0.346)	0.981
Security camera	$0.615^{(400)}$	(0.487)	$0.619^{(733)}$	(0.486)	$0.617^{(1,133)}$	(0.486)	0.619
Security guard	$0.255^{(1,100)}$	(0.436)	$0.262^{(1,433)}$	(0.440)	$0.259^{(2,533)}$	(0.438)	0.324
Local recipient	0.936	(0.245)	0.932	(0.252)	0.934	(0.249)	0.349
No English	0.324	(0.468)	0.315	(0.464)	0.319	(0.466)	0.215
Understood situation	$5.733^{(423)}$	(0.673)	$5.748^{(757)}$	(0.631)	$5.740^{(1,180)}$	(0.652)	0.559
Busy	$1.932^{(4)}$	(1.696)	$1.961^{(2)}$	(1.715)	$1.947^{(6)}$	(1.706)	0.333
Observations	7,890		8,214		16,104		

Notes: We coded "Age 40+" as 1 if the recipient was judged to be 40 years or older, and 0 otherwise. "Male" was coded as 1 if the recipient was male and 0 otherwise. "Computer" was coded as 1 if there was a computer present at the recipient's desk, and 0 otherwise. "Coworkers" was coded as 1 if a recipient's coworkers participated in the exchange, and 0 otherwise. "Other bystanders" was coded as 1 if any bystanders witness the exchange and 0 otherwise. The variables "Hotel," "Bank," "Cultural," "Public," and "Postal" represent the five types of institutions in which the experiments were performed. "Security camera" was coded as 1 if a security camera was visible during the exchange, and 0 otherwise. "Security guard" was coded as 1 if a security guard was present, and 0 otherwise. "Local recipient" was coded as 1 if the recipient was judged to be a local resident, and 0 otherwise. "No English" was coded as 1 if the recipient was not able to communicate with the research assistant in English, and 0 otherwise. "Understood situation" was a rating of whether the recipient understood the situation on a scale from 0 (*not at all*) to 6 (*fully understood*). "Busy" was a rating of how busy the recipient was when the wallet was turned in, on a scale from 0 (*not at all*) to 6 (*very busy*). The last column presents *P*-values for the null hypothesis of perfect randomization (χ^2-tests, except for "understood situation" and "busy" where we perform a Kruskal-Wallis *H* test). Not all variables were collected in each country and there are a few missing answers in the drop-off survey. Numbers in superscript parenthesis indicate the number of drop-offs that are missing in each condition.

Tab. S6. Analysis of rejections

	All countries (1)	UK, Poland, and US (2)
Money	0.182	−0.270
	(0.121)	(0.380)
BigMoney		0.617
		(0.559)
Money-NoKey		0.517
		(0.547)
Constant	0.130	0.380
	(0.161)	(0.427)
Controls:		
Institution FE	yes	yes
City FE	yes	yes
Money = BigMoney		0.077
Money = Money-NoKey		0.117
BigMoney = Money-NoKey		0.878
Wald test		0.184
Observations	16204	2959
Adjusted R^2	0.009	0.006

Notes: OLS estimates with robust standard errors in parentheses. Column 1 shows the results for treatment Money and NoMoney in all 40 countries. Column 2 shows the results for all four treatments in the United Kingdom, Poland, and the United States. The dependent variable is a dummy variable indicating whether the recipient refused to take the wallet. "Money," "BigMoney," and "Money-NoKey" are treatment indicators. The omitted category is the treatment "NoMoney." All models include city and institution fixed effects. The bottom of the table reports P-values from t-tests for equality of the treatment coefficients and a Wald test of the joint significance of all treatments. Significance levels: * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.

Tab. S7. Analysis of response times

	All countries	UK, Poland, and US
	(1)	(2)
Money	−0.082	−0.053
	(0.131)	(0.491)
BigMoney		−0.142
		(0.497)
Money-NoKey		−0.232
		(0.510)
Constant	0.760	−0.185
	(0.630)	(0.492)
Controls:		
Institution FE	Yes	Yes
City FE	Yes	Yes
Money = BigMoney		0.840
Money = Money-NoKey		0.666
BigMoney = Money-NoKey		0.838
Wald test		0.964
Observations	7340	1711
Adjusted R^2	0.015	−0.004

Notes: OLS estimates with robust standard errors in parentheses. The dependent variable is the response time in days. Column 1 shows the results for treatment Money and NoMoney in all 40 countries. Column 2 shows the results for all four treatments in the United Kingdom, Poland, and the United States. The dependent variable is the response time in days. "Money," "BigMoney," and "Money-NoKey" are treatment indicators. The omitted category is the treatment "NoMoney." All models include city and institution fixed effects. The bottom of the table reports P-values from t-tests for equality of the treatment coefficients and a Wald test of the joint significance of all treatments. Significance levels: * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.

Survey Experiments

We conducted nationally representative online survey experiments in the United Kingdom, Poland, and the United States to investigate self-reported motives for deciding to return or keep a lost wallet. We conducted surveys in the UK and US in English. For Poland, we hired two professional translators — one for the Polish translation and the other to translate it back to English. We did this to ensure that the meaning of the questions were not lost in translation.

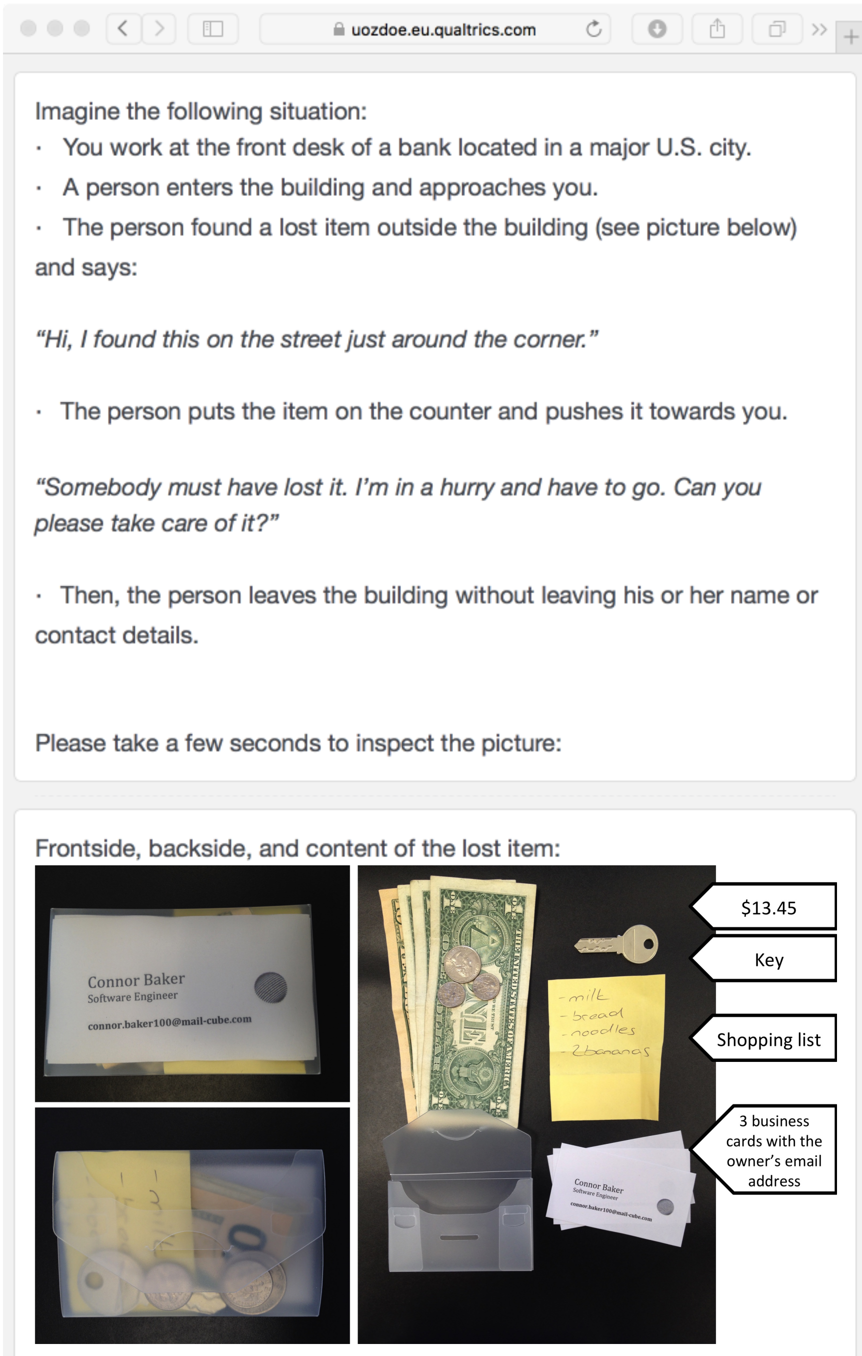
We sampled a total of 2,525 respondents through a Qualtrics online sample ($n = 829$ in the UK, $n = 809$ in Poland, and $n = 887$ in the US). To qualify for participation, individuals had to pass a simple attention check and meet the demographic quotas (based on age, gender, and residence) set by Qualtrics to construct the representative samples. Participants received a flat payment of US $4.00 for their participation.

We randomly assigned participants to one of our four treatments corresponding to the NoMoney, Money, BigMoney, and Money-NoKey condition. Participants were told the study was about lost and found property, and then asked to rate their knowledge of lost property laws. They then read a brief description of a typical drop-off scenario and viewed a picture of the wallet and its contents. The particular description and picture of the wallet varied according to the condition. We also randomized the owner's name and the type of institution. Fig. S3 provides an example of how this information was presented to participants.

After reading the scenario, participants completed several blocks of questions. In the first block, participants were asked how likely was it they would receive a financial reward from the owner if they were to contact him about the wallet, and responded on an 11-point scale from 0% to 100% in 10% increments. They were then asked, assuming the owner offered a financial reward, how much money they thought the owner would give. Participants provided their response in an open-text field.

In the second block, participants were asked the following questions on 11-point scales (0 = *not at all*, 10 = *very much*): "How concerned would you be with other people's impression of you

Fig. S3. Example survey scenario



Notes: Scenarios and pictures were adjusted according to experimental condition and country. We also randomly varied the owner’s name and type of institution in the scenario.

if you do not contact the owner?", "How important do you think is the lost item for its owner?", "To what extent would it feel like stealing if you do not contact the owner?", and "How concerned would you be that you get punished if you do not contact the owner?" The order of questions in this block were randomized for each participant.

In the third block, participants were asked to guess the annual income of the owner compared to the average person in their country on a seven-point scale (-3 = *much lower than the average person*, +3 = *much higher than the average person*). In the fourth block, participants were asked how likely they would be to contact the owner to return the lost item, and also how likely that someone else would contact the owner to return the lost item in such a situation. For both questions participants responded on an 11-point scale from 0% to 100% in 10% increments.

We then included a number of exploratory questions. Participants were asked if they personally have ever lost a wallet, a mobile phone, or a key, as well as if they have ever found a lost wallet, mobile phone, or key. For each item they responded either yes (coded as 1) or no (coded as 0). In another block participants completed seven items from the Empathic Concern subscale of the Interpersonal Reactivity Index (*48*). Participants also completed six items from the Impression Management subscale of the Balanced Inventory of Desirable Responding (*49*), and a four-item measure of general attitudes about honesty. Lastly, participants were presented with several misbehaviors (e.g., cheating on one's taxes) and asked to assess the degree that most other people would consider the behavior appropriate or inappropriate on a four-point scale (-2 = *very inappropriate*, +2 = *very appropriate*).

As an additional attention check, we asked participants to recall key details from the study. We first asked them to list the contents of the wallet in a series of open-text boxes. We then asked them to identify the name of the owner from a list of 6 options. Finally, we asked them to recall the amount of money in the wallet in an open-text box.

Prediction Study: Non-expert Sample

We conducted an online survey in the United States to investigate lay beliefs about the relationship between civic honesty and monetary incentives. Our sample consisted of 299 U.S. adults from Amazon.com's Mechanical Turk labor market (58% male, 42% female; *M* age = 35.49, *SD* = 10.66). To qualify for participation, individuals had to take the survey using a non-mobile device (such as a desktop or laptop computer) and pass a simple attention check. Participants received a flat payment of US $0.50 for their participation, along with the opportunity to win a $5.00 bonus.

Participants were told that we had recently conducted a study in 25 US cities, and their job was to predict the outcomes of the study. We first described the general design of our lost wallet experiments, then provided participants with details about the exact procedure, the wallets we turned in, and details about three of our experimental conditions (NoMoney, Money, and BigMoney). Participants were also provided with an image of the wallets similar to that in Fig S3. We then asked participants to predict reporting rates (from 0-100%) for each condition. We informed participants that they should try their best to be accurate, as the most accurate 5% of participants in the study would receive a bonus payment of $5.00. All responses were made on the same page using slider scales from 0 to 100.

On the next page we probed participants' beliefs about the relevant motivations operating in each of our experimental conditions. We first asked participants to consider the following three issues that our recipients may have been considering when deciding to return or not return a wallet: (i) how tempted would the recipient be to keep the money in the wallet, (ii) how concerned would the recipient be for the owner, and (iii) how much would the recipient feel like they were a thief if they did not return the wallet. Participants estimated the relative importance of these three concerns for each condition on 100-point slider scales, with higher numbers indicated greater importance. For each condition, responses for the three concerns were required to sum to 100.

Afterwards, participants provided basic demographic information including their age, gender, ethnicity, educational level, employment status, and household income.

Prediction Study: Expert Sample

We conducted a follow-up online survey to investigate expert predictions about the relationship between civic honesty and monetary incentives. To do so, we surveyed a group of academic economists whose email addresses were publicly available on the Research Papers in Economics repository website (`http://repec.org`).

We culled email addresses for economists who have published in the last five years, and who ranked in the top 5% in at least one of the following dimensions on the website: "average rank," "citations," "citations, discounted by age," "h-index," "abstract views," and "downloads." To exclude economists who were likely to be familiar with our project, we excluded anyone from our email list who was affiliated with a research institute in Zurich or on the website's expert lists for experimental economics, cognitive and behavioral economics, norms and social capital, or prospect theory. This procedure yielded 2,283 email addresses. We sent out an invitation to participate in the study, and received 294 completed responses. For our analysis we excluded 15 respondents who reported familiarity with our lost wallet experiments, yielding a final sample of 279 participants (88% male, 12% female; *M* age = 54.60, *SD* = 11.64). The overwhelming majority of respondents were university professors (95%), with 71% at the rank of full professor.

Participants were given the same instructions and were asked to make the same predictions as in our previous prediction study, but were not asked to complete the motivation items on self-interest, altruism, and theft aversion. Participants were informed up front that the three most accurate respondents would receive a US $100 bonus which they could keep or donate to charity. At the end of the survey we asked respondents to report their gender, age, current academic status/ranking, and whether they were previously familiar with our lost wallet experiments.

Supplementary Text 1: A Conceptual Framework for Civic Honesty

We model a recipient's decision to return a lost wallet as follows. A recipient chooses an action $a \in \{0,1\}$ to either keep the wallet ($a = 0$) or return the wallet along with its content ($a = 1$). The recipient's decision is determined by four factors. The first factor reflects the effort necessary to return the wallet. The recipient incurs an effort cost c when returning the wallet, such as the time required to contact the owner. The second factor reflects the potential material benefits to the recipient. If the recipient decides to keep the wallet, then her material payoff increases by the amount of money m in the wallet. The third factor reflects potential altruistic concerns from the recipient towards the owner, captured by the weight α that the recipient places on the potential externality. If the recipient fails to return the wallet then she can internalize the costs to the owner, which includes the money inside the wallet (m) along with anything else inside the wallet thought to be valuable to the owner (v). Based on prior empirical work (*14, 50, 51*), we assume that the recipient cannot value the wallet more than its owner ($0 \le \alpha < 1$). The fourth factor reflects self-image concerns, captured by the weight γ (hereafter what we call "theft aversion"). If the recipient fails to return the wallet then she may consume a negative self-image resulting from thinking of herself as a dishonest person. The weight placed on theft aversion is assumed to be non-negative, $\gamma \ge 0$. Based on these four factors, an individual chooses action a in order to maximize the following objective function:

$$\max_{a\in\{0,1\}} \{(1-a)m + a\alpha(m+v) - (1-a)\gamma m - ac\}. \tag{2}$$

As is clear from equation (2), we assume the non-pecuniary costs of failing to return the wallet (captured by α and γ) increase linearly with the amount of money inside the wallet.[6] It follows

[6] This is a reduced form representation consistent with signaling models such as (*52*), where recipients are concerned about their social or self-image. Returning a wallet with greater amounts of money is a costlier signal about the recipient's honesty and therefore yields a higher reputational benefit than a wallet with smaller amounts of cash.

from equation (2) that a recipient will return a wallet if and only if

$$\alpha v + m(\alpha + \gamma - 1) \geq c. \tag{3}$$

Note that in our framework theft aversion depends on the amount of money in the wallet, whereas altruistic concerns for the owner depend on the amount of money as well other contents in the wallet thought to be valuable to the owner. Accordingly, recipients sufficiently high in altruism (i.e., a high α) would be compelled to return the wallet even when it contains little or no money. By contrast, recipients who are theft averse (i.e., high γ) would be compelled to return a wallet only when it contains sufficiently large amounts of money.

When we apply the framework to our current experiments, we obtain four potential types of recipients. The first type involves recipients primarily motivated by material self-interest (i.e., low α and low γ), who will never return a wallet regardless of its contents. Our second type involves recipients who are sufficiently altruistic and theft averse (i.e., high α and high γ) who will always return the wallet regardless of its contents (so long as such concerns outweigh the effort costs of returning the wallet).

The third and fourth types are unique in that their behavior will depend on the wallet's contents. Our third type involves recipients high in altruism but low in theft aversion (i.e., high α and low γ), who will return a wallet with little to no money but will keep a wallet when it contains sufficiently large amounts of cash. These individuals are primarily motivated by altruistic concerns for low amounts of money, but self-interest dominates for larger amounts of money. Formally this type is comprised of individuals where

$$\frac{c + m(1-\gamma)}{v+m} > \alpha > \frac{c}{v}. \tag{4}$$

Our fourth type involves recipients low in altruism but high in theft aversion (i.e., low α and high γ), who will fail to return a wallet with little to no money but return a wallet when it contains larger amounts of cash. These individuals will not be sufficiently motivated by concern for the owner's

Psychological costs could also be represented in other forms, such as negative emotional costs (*53, 54*) or a desire to adhere to social norms (*55, 56*).

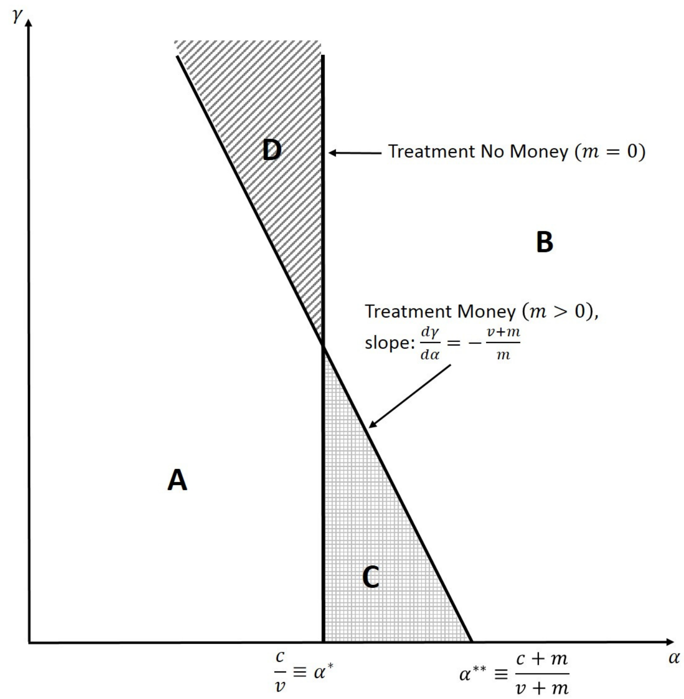
welfare to return wallets with relatively small amounts of money, but theft aversion concerns will dominate for larger amounts of money. Formally this type is comprised of individuals where

$$\frac{c}{v} > \alpha > \frac{c+m(1-\gamma)}{v+m}. \tag{5}$$

The distribution of types in the population determines the nature of the relationship between the reporting rate and the amount of money in the wallet. Figure S1 illustrates this dynamic along a α/γ-plane for the NoMoney and Money conditions. In the NoMoney condition, recipients with sufficiently high altruistic concerns ($\alpha > \alpha^*$) will return the wallet, while all other recipients will not; the separation between these two response types is denoted by the vertical line in Figure S1. In the Money condition, recipient types are distinguished by the line with slope $-(v+m)/m$ which intersects the α-axis to the right of α^* at $\alpha^{**} = (c+m)/(v+m)$. The two lines divide the plane into four regions. Recipients in region A fail to return the wallet in both conditions because they are self-regarding, reflecting our first type (low α and low γ). Recipients in region B will return the wallet in both treatments because they are sufficiently altruistic and theft averse, reflecting our second type (high α and high γ). Region C consists of recipients who are altruistic enough to return the wallet in the NoMoney condition but fail to return the wallet in the Money condition due to self-interest, reflecting our third type (i.e., high α and low γ). Finally, region D consists of recipients who do not reach the threshold of altruism α^* in the NoMoney condition and therefore do not return the wallet, but who are sufficiently motivated by theft aversion to return the wallet in the Money condition.

Based on our framework, treatment differences in reporting rates reflect the distribution of types in the population. The fact that reporting rates are relatively higher in the Money condition suggest that recipient types in region D are more prevalent than those in region C. An analogous line of reasoning can be applied to explain the increase in civic honesty in the BigMoney condition relative to the NoMoney and Money conditions.

Fig. S4. Response patterns as a function of altruism (α) and theft aversion (γ)



This figure illustrates response patterns for each of four behavioral types as a function of altruism (α) and theft aversion (γ). Recipients in region A will not report a wallet in either treatment. In contrast, recipients in region B will always report a wallet, regardless of whether it contains money or not. Recipients in region C are sufficiently altruistic to return a wallet in the NoMoney condition, but their degree of theft aversion is not large enough to compensate the temptation to pocket the money in Money condition. Finally, region D comprises recipients who are not sufficiently altruistic to report a wallet with no money, but their degree of theft aversion is strong enough to induce them to return the wallet in the Money condition.

Supplementary Text 2: Results

Behavioral Data from Lost Wallet Experiments

Civic Honesty Across Countries We first examine reporting rates in the NoMoney and Money conditions for all 40 countries. Overall, 51% of recipients in the Money condition reported the wallet compared to 40% of recipients in the NoMoney condition ($z = 14.18$, $P < 0.0001$). We observe an increase in reporting rates for the Money condition relative to the NoMoney condition in 38 out of 40 countries, and this effect is statistically significant at the 5% level for 19 countries after adjusting for the pairwise comparison false discovery rate (*26*). Furthermore, in neither of the two countries that displayed a reduction in reporting rates in the Money condition was the decline statistically significant ($z = 1.47$, $P = 0.141$ for Mexico; $z = 0.19$, $P = 0.853$ for Peru).

Table S8 displays the results when aggregated across all 40 countries. For the table as well as all subsequent analyses, we use ordinary least squares (OLS) regression with robust standard errors. Responses are coded as 100 if the wallet was reported and 0 otherwise. We use OLS for purposes of simplicity and clarity because coefficients can be directly interpreted as percentage point changes; using nonlinear models such as logistic regression return virtually identical results. Column 1 of Table S8 indicates that reporting rates increase by 10.8 percentage points in the Money relative to the NoMoney condition when including city, institution, and treatment fixed effects[7] ($t_{16941} = 15.16$, $P < 0.001$).

Column 2 of Table S8 indicates that our treatment effect holds when also controlling for additional recipient and situational characteristics. This specification also finds that these additional characteristics also influenced reporting rates independent of our experimental conditions. On average men were roughly 2 percentage points less likely than women to report a wallet ($t_{16928} = 2.78$, $P = 0.005$), and older recipients (i.e., those judged to 40 years or older) were 2 percentage points less likely to report a wallet ($t_{16928} = 2.75$, $P = 0.006$). The presence of a computer at the re-

[7]Controlling for the other two experimental conditions does not affect estimates of the Money coefficient, but provides added precision when estimating our other control variables.

Tab. S8. Reporting rates in the Money and NoMoney condition

	(1)	(2)
Money	10.828***	10.792***
	(0.714)	(0.712)
Male		−2.076**
		(0.747)
Age 40+		−2.030**
		(0.738)
Computer		6.874***
		(0.969)
Coworkers		4.675***
		(0.765)
Other bystanders		−3.900***
		(0.795)
Constant	34.620**	33.302**
	(11.434)	(11.112)
Controls:		
Institution FE	yes	yes
City FE	yes	yes
Treatments	yes	yes
Observations	17303	17295
Adjusted R^2	0.178	0.185

Notes: OLS estimates with robust standard errors in parentheses. The dependent variable in all models takes on the value 100 if a wallet was reported and 0 otherwise. "Money" is a dummy for treatment Money (we also include an indicator for treatments "BigMoney" and "Money-NoKey"). The omitted category in this table is the treatment "NoMoney." All models further include city and institution fixed effects. In column 2, we also include binary control variables for individual and situational factors, including a recipient's age (above 40 years) and gender (male), as well as the presence of a computer, coworkers, and other bystanders. Significance levels: * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.

cipient's workstation increased the likelihood of reporting a wallet ($t_{16928} = 7.10$, $P < 0.001$), as did the presence of other coworkers ($t_{16928} = 6.11$, $P < 0.001$). The latter of the two findings is unsurprising given that, in addition to the possibility of increased social monitoring, the presence of other coworkers may have also reduced recipients' workload. By contrast, the presence of other bystanders (excluding coworkers) decreased reporting rates ($t_{16928} = 4.90$, $P < 0.001$). One possibility for this result is that the increase in workload by having bystanders present exerted a larger influence on recipients' behavior than did the additional social pressure brought about by the bystander's presence.

Civic Honesty under High Stakes We next examine reporting rates for the three countries in which we conducted the BigMoney condition alongside our Money and NoMoney conditions (N = 2,932). Despite the higher incentive to steal, recipients were more likely to report a lost wallet when it contained greater amounts of money. Across the three countries, 46% of the recipients reported the wallet in the NoMoney condition, which increased to 61% in the Money condition and increased even further to 72% in the BigMoney condition ($z > 4.40$ for all pairwise comparisons, $P < 0.001$). Column 1 in Table S9 shows that, when controlling for situational and recipient characteristics, the average share of recipients who reports a wallet increases by almost 16 percentage points in the Money relative to the NoMoney condition ($t_{2846} = 6.73$, $P < 0.001$). The BigMoney condition increases the reporting rate by 25 percentage points, on average, relative to the NoMoney condition ($t_{2846} = 9.86$, $P < 0.001$), and the difference between the BigMoney and Money conditions is also significant ($t_{2846} = 3.92$, $P < 0.001$). Columns 2-4 show that the increasing trend in civic honesty for larger monetary stakes holds for all three countries.

Testing for Altruism To examine the role of altruism, we compare the Money condition to the Money-NoKey condition for the three countries where we conducted both treatments ($N = 2,932$). Wallets from these two conditions contain the same contents with the exception of the key, which

Tab. S9. Reporting rates in NoMoney, Money, and BigMoney condition

	UK, Poland, and US	United Kingdom	Poland	United States
	(1)	(2)	(3)	(4)
Money	15.940***	23.106***	3.310	18.301***
	(2.370)	(3.851)	(4.690)	(3.934)
BigMoney	25.235***	35.941***	11.761**	27.832***
	(2.558)	(4.567)	(4.410)	(4.260)
Constant	35.506***	25.763**	59.380***	34.445**
	(8.517)	(9.345)	(11.216)	(11.291)
Controls:				
Recipient	yes	yes	yes	yes
Situation	yes	yes	yes	yes
Institution FE	yes	yes	yes	yes
City FE	yes	yes	yes	yes
Other treatments	yes	yes	yes	yes
Money = BigMoney	0.000	0.001	0.058	0.027
Observations	2926	1132	794	1000
Adjusted R^2	0.091	0.122	0.050	0.100

Notes: OLS estimates with robust standard errors in parentheses. Column 1 presents the results for all three countries, column 2 for the United Kingdom, column 3 for Poland, and column 4 for the United States. The dependent variable in all models takes on the value 100 if a wallet is reported and 0 otherwise. “Money” and “BigMoney” are treatment indicators (we also include an indicator for our “Money-NoKey” treatment but report those estimates in Table S10). The omitted category in this table is the treatment “NoMoney.” All models include binary control variables for recipient and situational characteristics, including a recipient’s age (above 40 years) and gender (male), as well as the presence of a computer, other people, and coworkers. All models include city and institution fixed effects. The “Money = BigMoney” row reports P-values from t-tests for equality of the treatment coefficients. Significance levels: * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.

is valuable to the owner of the wallet but not to the recipient.[8] As a result, altruistic concerns should be responsible for any differences in reporting rates between the Money and Money-NoKey conditions. Shown in Table S10, we do find relatively fewer wallets were reported when they did not contain a key. Column 1 indicates that the average reporting rate across countries decreased by more than 9 percentage points in the Money-NoKey condition relative to the Money condition ($t_{2846} = 3.70$, $P < 0.001$). Columns 2-4 show that this pattern holds for all three countries, though the difference was statistically significant only for the UK and Poland (12 and 10 percentage points, respectively).

[8] In the representative survey experiments, we asked participants to evaluate the importance of the wallet to the owner on a 11-point scale from not at all (0) to very much (10). Consequently, respondents tended to recognize the value of the key to the owner. On average, respondents considered the wallet in the Money-NoKey condition to be 2.32 points (or 0.86 standard deviations) less important to the owner compared than the wallet in the Money condition ($t_{1120} = 14.33$, $P < 0.001$). This comparison is in the same direction and statistically significant when examining each country separately (all P-values < 0.001).

Tab. S10. Reporting rates in Money-NoKey condition

	UK, Poland, and US	United Kingdom	Poland	United States
	(1)	(2)	(3)	(4)
Money-NoKey	-9.185^{***}	-11.750^{**}	-9.820^{*}	-2.927
	(2.482)	(3.832)	(4.743)	(4.433)
Constant	51.446^{***}	48.869^{***}	62.690^{***}	52.746^{***}
	(8.393)	(8.971)	(11.068)	(11.373)
Controls:				
Recipient	yes	yes	yes	yes
Situation	yes	yes	yes	yes
Institution FE	yes	yes	yes	yes
City FE	yes	yes	yes	yes
Other treatments	yes	yes	yes	yes
Observations	2926	1132	794	1000
Adjusted R^2	0.091	0.122	0.050	0.100

Notes: OLS estimates with robust standard errors in parentheses. Column 1 presents the results for all three countries, column 2 for the United Kingdom, column 3 for Poland, and column 4 for the United States. The dependent variable in all models takes on the value 100 if a wallet is reported and 0 otherwise. "Money-NoKey" is a treatment indicator (we also include indicators for treatments "NoMoney" and "BigMoney" but do not report their estimates for ease of exposition). The omitted category in this table is the treatment "Money." All models include binary control variables for individual characteristics and situational factors, including a recipient's age (above 40 years) and gender (male), as well as the presence of a computer, other people, and coworkers. All models include city and institution fixed effects. Significance levels: * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.

Survey Data from Nationally Representative Samples

Since we collected survey data to measure how possible psychological motives to report a lost wallet differ according to wallet content, we restrict our analysis to participants who were able to correctly recall the amount of money inside the wallet described to them (rounded to the nearest integer). This leaves us with a sample of 2,160 participants from our original sample of 2,525. When we do not exclude any participants we find largely similar results (displayed in Table S12) to those reported below.

Evidence for Theft Aversion In our survey experiments, we asked participants to rate the extent to which failing to report a wallet felt like stealing. Column 1 in Table S11 shows that across the three countries, respondents reported that failing to return a wallet would feel more like stealing when the wallet contained greater amounts of money. Relative to the NoMoney condition, the average score increased by 1.57 points (or 0.47 standard deviations) in the Money condition, and by 2.08 points (or 0.64 standard deviations) in the BigMoney condition ($t_{2150} = 7.72$, $P < 0.001$ for Money; $t_{2150} = 10.41$, $P < 0.001$ for BigMoney). The difference between the Money and BigMoney condition was also significant ($t_{2150} = 2.71$, $P = 0.007$). In contrast, we failed to observe a reliable difference in responses between the Money and Money-NoKey conditions ($t_{2150} = 1.13$, $P = 0.259$). This suggests that anticipated costs due to theft aversion depend on the amount of money in the wallet, but do not meaningfully depend on other contents that are only valuable to the owner.

In the survey we also asked respondents to report the likelihood they would contact the owner to return the wallet (from 0-100%). Naturally such self-reports should be interpreted with caution, and indeed we find responses were overly optimistic when compared with the behavioral data (average estimates ranged between 88% and 93% across countries). Nevertheless, the pattern of treatment differences in self-reported likelihood of returning wallet follow the same rank-ordering as those from our lost wallet experiments (see column 2 in Table S11), and so we use our self-report data as a proxy for exploring possible motives for returning a lost wallet.

Tab. S11. Survey responses across experimental conditions

	Theft aversion concerns	Stated likelihood of reporting (in %)		
	(1)	(2)	(3)	(4)
Money	1.570***	2.400*	−0.748	−0.368
	(0.203)	(0.985)	(0.966)	(0.941)
BigMoney	2.076***	3.847***	−0.315	−0.928
	(0.200)	(0.975)	(0.989)	(0.979)
Money-NoKey	1.358***	−2.454*	−5.177***	−1.843
	(0.201)	(1.171)	(1.131)	(1.163)
Theft aversion concerns			2.005***	1.690***
			(0.161)	(0.152)
Perceived importance to owner				1.283***
				(0.180)
Fear of punishment				0.133
				(0.106)
Constant	6.512***	86.414***	73.357***	65.609***
	(0.224)	(1.217)	(1.675)	(2.126)
Controls:				
Institution FE	yes	yes	yes	yes
Country FE	yes	yes	yes	yes
Money = BigMoney	0.007	0.120	0.623	0.516
Money = Money-NoKey	0.259	0.000	0.000	0.165
Observations	2160	2160	2160	2160
Adjusted R^2	0.053	0.029	0.159	0.188

Notes: OLS estimates with robust standard errors in parentheses. In column 1 the dependent variable is our proxy for theft aversion concerns which is measured by the question "To what extent would it feel like stealing if you do not contact the owner?" with possible answers ranging from "not at all" (0) to "very much" (10). The dependent variable in columns 2-4 is the likelihood that participants would report the wallet (as a percentage). "Money," "BigMoney," and "Money-NoKey" are treatment indicators. All models include country and institution fixed effects. The bottom of the table reports P-values from t-tests for equality of the treatment coefficients. Significance levels: * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.

Tab. S12. Survey responses across experimental conditions, full sample

	Theft aversion concerns	Stated likelihood of reporting (in %)		
	(1)	(2)	(3)	(4)
Money	1.501***	2.919**	−0.121	0.219
	(0.189)	(0.958)	(0.939)	(0.916)
BigMoney	1.742***	3.708***	0.181	−0.423
	(0.190)	(0.968)	(0.975)	(0.960)
Money-NoKey	1.225***	−1.736	−4.217***	−0.786
	(0.189)	(1.124)	(1.082)	(1.112)
Theft aversion concerns			2.025***	1.699***
			(0.151)	(0.140)
Perceived importance to owner				1.436***
				(0.172)
Fear of punishment				0.059
				(0.099)
Constant	6.653***	85.616***	72.143***	63.917***
	(0.208)	(1.158)	(1.573)	(2.003)
Controls:				
Institution FE	yes	yes	yes	yes
Country FE	yes	yes	yes	yes
Money = BigMoney	0.167	0.371	0.716	0.427
Money = Money-NoKey	0.110	0.000	0.000	0.312
Observations	2525	2525	2525	2525
Adjusted R^2	0.039	0.023	0.152	0.185

Notes: OLS estimates with robust standard errors in parentheses. Full sample, including participants that failed our recall attention check. In column 1 the dependent variable is our proxy for theft aversion concerns measured by the question “To what extent would it feel like stealing if you do not contact the owner?” with possible answers ranging from “not at all” (0) to “very much” (10). The dependent variable in columns 2-4 is the likelihood that participants would report the wallet (as a percentage). “Money,” “BigMoney,” and “Money-NoKey” are treatment indicators. All models include country and institution fixed effects. The bottom of the table reports P-values from t-tests for equality of the treatment coefficients. Significance levels: * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.

Column 3 in Table S11 shows that theft aversion concerns were positively related to one's stated likelihood of reporting a wallet ($t_{2149} = 12.44$, $P < 0.001$). Furthermore, compared to the model displayed in column 2 that does not control for theft aversion, the model in column 3 provides a substantially better fit to the data (adjusted R-squared increases from 0.029 to 0.159) and the coefficients for the Money and BigMoney conditions shrink and are no longer statistically significant. To the extent such self-reports extend to real behavior, theft aversion may partly explain why people are more likely to return a lost wallet with greater amounts of money inside. Finally, column 4 also includes a measure of perceived importance of the wallet to the owner, which serves as a proxy for altruistic concerns, and a measure for the subjective fear of punishment if the wallet is not reported. We find that both the perceived importance of the wallet and the aversion to viewing oneself as a thief are positively related to the stated likelihood of reporting the wallet ($t_{2147} = 7.14$, $P < 0.001$ and $t_{2147} = 11.11$, $P < 0.001$, respectively). This suggests that both altruism and theft aversion concerns are relevant to reporting a lost wallet, and that the two operate independently of each other. In contrast, self-reported fear of punishment is not significantly correlated with the stated likelihood of reporting the wallet ($t_{2147} = 1.26$, $P = 0.208$). Thus, if anything, threat of punishment plays only a minor role in reporting a lost wallet.

The pattern of results displayed in Table S11 suggests that theft aversion explains why the reporting rate increases with the amount of money in the wallet, but not with the presence or absence of the key. To test this hypothesis, we conducted a series of mediation analyses. For the first mediation test, we restricted observations to the three conditions that only varied the amount of money in the wallet (NoMoney, Money, and BigMoney conditions). Using the NoMoney condition as our reference variable, we calculated indirect paths {Money $\rightarrow$ theft aversion $\rightarrow$ Likelihood of reporting} and {BigMoney $\rightarrow$ theft aversion $\rightarrow$ Likelihood of reporting} using bootstrapped standard errors with 10,000 resamples (*57*). Consistent with our hypothesis, theft aversion mediated the relationship between the amount of money inside the wallet and the likelihood of reporting a lost wallet (indirect $b_{Money} = 2.47$, SE = 0.42, $P < 0.001$; indirect $b_{BigMoney} = 3.26$, SE = 0.48, $P < 0.001$). Furthermore, the direct effect in both conditions was nonsignificant after accounting for the indirect effect of theft aversion (direct $b_{Money} = 0.003$, SE = 0.97, $P = 0.997$; direct $b_{BigMoney} = 0.43$, SE = 0.99, $P = 0.666$).

We then conducted a second mediation test based on our framework's assumption that altruism, rather than theft aversion, should explain the difference in reporting rates between the Money and Money-NoKey conditions. Restricting observations to only those two conditions, we conducted a similar path analysis as before except this time for the indirect paths {Money-NoKey $\rightarrow$ Perceived harm to owner $\rightarrow$ Likelihood of reporting} and {Money-NoKey $\rightarrow$ Theft aversion $\rightarrow$ Likelihood of reporting}. Consistent with our conceptual framework, we find that our proxy for altruistic concerns (perceived harm to the owner) reliably mediates the difference between the two conditions (indirect $b = -3.58$, SE = 0.57, $P < 0.001$) while theft aversion does not (indirect $b = -0.45$, SE = 0.41, $P = 0.274$). Furthermore, the direct effect of experimental condition was nonsignificant after accounting for our indirect effects (direct $b = -0.96$, SE = 1.10, $P = 0.381$). Taken together these results are consistent with the hypothesis that theft aversion explain why the reporting rate increases with the amount of money in the wallet, but does not explain why the reporting rate decreases with the absence the key.

Prediction Data: Non-expert Sample

We examined whether people anticipate our behavioral results by asking online participants to predict reporting rates in the US for wallets that contained \$0, \$13.45, and \$94.15. Contrary to the behavioral data, respondents predicted that reporting would be highest when the wallet contained no money ($M = 72.71$, $SD = 29.47$), lower when the wallet contained a modest amount of money ($M = 65.04$, $SD = 24.01$), and lower still when the wallet contained a substantial amount of money ($M = 54.55$, $SD = 28.88$). All three predictions were reliably different from one another (Table S13, Column 1; $t_{298} \geq 6.40$, $P < 0.001$ for all pairwise comparisons). For each condition we also compared the average predicted change to the actual change in reporting rates. The predicted change in reporting rates was always lower (i.e., more cynical) than the actual change in reporting rates ($t_{298} \geq 12.16$, $P < 0.001$ for all pairwise comparisons).

We next examined response profiles within participants.[9] As the amount of money inside the wallet increased, 64% predicted a monotonic decrease in civic honesty, 18% predicted a monotonic increase in civic honesty, 3% predicted no change, and 15% displayed non-monotonic predictions. Using a sign test (coded as –1 = predicted a decrease in civic honesty, +1 = predicted an increase in civic honesty, 0 = all remaining responses), we find that reliably more participants expected rates of civic honesty to decrease than increase as wallet amounts became larger ($P < 0.001$).

Participants also reported their beliefs about the relative share of different motivations operating in each condition. On average participants expected self-interest to grow and altruistic concerns to shrink for wallets containing relatively more money. Compared to the NoMoney condition, participants expected the temptation of recipients to pocket the money to increase by 18.95 points (or 0.93 standard deviations) in the Money condition, and by 36.98 points (or 1.26 standard deviations) in the BigMoney condition ($t_{298} > 16.00$, $P < 0.001$ for both comparisons). The difference between the Money and BigMoney conditions was also significant ($t_{298} = 14.15$, $P < 0.001$). We see the reverse pattern for beliefs about altruistic concerns by recipients towards the owner of the

[9] We assume weak monotonicity when calculating percentages for response profiles. Results from our sign-tests do not meaningfully change when response profiles are instead calculated assuming strong monotonicity.

wallet. Relative to the NoMoney condition, participants expected altruistic concerns to decrease by 23.95 points (or 0.92 standard deviations) in the Money condition, and by 42.15 points (or 1.31 standard deviations) in the BigMoney condition ($t_{298} > 15.80$, $P < 0.001$ for both comparisons). The difference between the Money and BigMoney conditions was also significant ($t_{298} = 14.96$, $P < 0.001$).

Recall that in the behavioral and self-report data, theft aversion appeared to play an important role in explaining variation across conditions appears. Respondents in our prediction study, on the other hand, afforded considerably less importance to concerns of theft aversion. Relative to the NoMoney condition, participants did expect concerns about viewing oneself as a thief to increase by 5 points (or 0.26 standard deviations) in the Money condition, and by 5.17 points (or 0.21 standard deviations) in the BigMoney condition ($t_{298} > 3.60$, $P < 0.001$ for both comparisons). The difference between the Money and BigMoney conditions was not statistically reliable ($t_{298} = 0.16$, $P = 0.875$). We also note differences in predicted theft aversion concerns across conditions were considerably smaller than those observed for predicted self-interest or altruism.

Lastly, we examined how inferences about motivations related to predictions of rates of civic honesty (columns 2–4, Table S13). Self-interest scores were inversely related to predicted reporting rates (column 2; $t_{298} = 9.54$, $P < 0.001$), and altruism scores were positively related to predicted reporting rates (column 3; $t_{298} = 6.53$, $P < 0.001$). In both cases, the adjusted R-squared increases by more than a factor of 2 relative to our baseline model in column 1, and the coefficients for our treatment coefficients shrink and are no longer statistically significant. However, as displayed in column 4, theft aversion concerns were not reliably associated with predicted reporting rates ($t_{298} = 1.36$, $P = 0.174$). When compared to our baseline model, including theft aversion concerns in the model does not meaningfully increase explained variance and our treatment coefficients do not decrease in size.

The pattern of results displayed in Table S13 suggest that respondents' inferences about self-interest and altruism, but not concerns of theft aversion, underly their beliefs that response rates will decline for wallets with relatively more money. To test this hypothesis, we conducted a series

Tab. S13. Predictions of reporting rates across experimental conditions

	(1)	(2)	(3)	(4)
Money	-7.672^{***}	1.613	0.198	-8.120^{***}
	(1.199)	(1.634)	(1.729)	(1.272)
BigMoney	-18.164^{***}	-0.041	-4.313	-18.627^{***}
	(2.297)	(2.893)	(2.965)	(2.300)
Self-interest		-0.490^{***}		
		(0.051)		
Altruism			0.329^{***}	
			(0.050)	
Theft aversion concerns				0.090
				(0.066)
Constant	72.709^{***}	77.533^{***}	49.534^{***}	70.950^{***}
	(1.706)	(1.716)	(3.820)	(2.180)
Money = BigMoney	0.000	0.327	0.007	0.000
Observations	299	299	299	299
Adjusted R^2	0.066	0.191	0.138	0.068

OLS estimates with participant-clustered standard errors in parentheses. The dependent variable is predicted reporting rates by recipients (from 0-100%). "Money" and "BigMoney" are treatment indicators. The omitted category is "NoMoney." The bottom of the table reports P-values from t-tests for equality of the treatment coefficients. Significance levels: * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.

of mediation analyses. For each our three motivation items,[10] we calculated the indirect pathway {Experimental conditions → Inferred motivation → Predicted reporting rate} using bootstrapped participant-clustered standard errors with 10,000 resamples (*57*). Consistent with the pattern suggested in Table S13, inferences of increasing self-interest and declining altruism each statistically mediate the relationship between experimental conditions and predicted reporting rates (self-interest results: indirect $b_{Money} = -9.29$, SE = 1.18, $P < 0.001$; indirect $b_{BigMoney} = -18.12$, SE = 2.18, $P < 0.001$; altruism results: indirect $b_{Money} = -7.87$, SE = 1.29, $P < 0.001$; indirect $b_{BigMoney} = -13.85$, SE = 2.20, $P < 0.001$). However, we fail to observe a reliable indirect effect of inferred theft aversion concerns on predicted reporting rates (indirect $b_{Money} = 0.45$, SE = 0.36, $P = 0.209$; indirect $b_{BigMoney} = 0.46$, SE = 0.37, $P = 0.216$). Thus, participants appeared to weight the role of self-interest and declining altruism, but not inferences of theft aversion, in predicting rates of civic honesty.

[10]We conducted separate mediation analyses for each motivation item rather than conduct a simultaneous mediation test for all three items, as the latter analysis would require us to remove at least one item due to collinearity (since inferences for the three items were required to sum to 100).

Prediction Data: Expert Sample

The results we observe from our expert sample were qualitatively similar to those from our MTurk sample, but considerably weaker in magnitude. On average, respondents predicted that reporting rates would be highest in the NoMoney condition ($M = 69.38$, $SD = 25.43$), followed by the Money condition ($M = 68.98$, $SD = 21.36$), and lowest in the BigMoney condition ($M = 65.70$, $SD = 23.15$). We compared conditions using an OLS regression with participant-clustered standard errors. Predicted reporting rates in the BigMoney condition were reliably lower than those in the NoMoney condition ($t_{278} = 2.05$, $P = 0.042$) and Money condition ($t_{278} = 2.50$, $P = 0.013$), but predicted reporting rates in the NoMoney and Money conditions did not reliably differ from one another ($t_{278} = 0.44$, $P = 0.660$). For each condition we also compared the average predicted change to the actual change in reporting rates. The predicted change in reporting rates was always lower (i.e., more cynical) than the actual change in reporting rates ($t_{278} \geq 8.70$, $P < 0.001$ for all pairwise comparisons).

We next examined response profiles within participants. As the amount of money inside the wallet increased, 49% predicted a monotonic decrease in civic honesty, 29% predicted a monotonic increase in civic honesty, 6% predicted no change, and 16% displayed non-monotonic predictions. Using a sign test (coded as –1 = predicted a decrease in civic honesty, +1 = predicted an increase in civic honesty, 0 = all remaining responses), we find that reliably more participants expected rates of civic honesty to decrease than increase as wallet amounts became larger ($P < 0.001$). In summary, experts in our sample held inaccurate beliefs, but to a lesser degree than our sample of MTurkers.

Cross-country Correlates of Civic Honesty

In this section, we explore possible explanations for cross-country differences in civic honesty. To address potential issues related to reverse causality, we primarily consider "deep" and historical explanatory variables which are plausibly exogenous to honest behavior and are considered formative to the development of society (*43*). To illustrate the value of this approach, consider that reporting rates in our study are positively correlated with contemporaneous measures of wealth (such as per capita income). From this correlation it is unclear whether country wealth leads to greater civic honesty or vice versa (or alternatively, some unobserved third variable influences both wealth and civic honesty). Now consider that, instead of wealth, we observed a correlation between a country's geographic terrain and civic honesty. Country terrain can be considered a deep variable because civic honesty is unlikely to influence geography, but geography could potentially influence civic honesty (by shaping citizen's interactions in ways that benefit or hinder cooperation). For this reason, using deep and historical variables is potentially more informative in explaining cross-country differences in civic honesty. We then extend our analysis to explore the role of culture and institutions, with the caveat that those factors may be endogenous[11] (*25, 58–60*).

We conducted a series of OLS regressions in which we regressed a given country-level variable onto individual decisions to report a wallet (for a full list of variables, see the "Country-level Correlates of Civic Honesty" subsection of Materials and Methods). As the rank-ordering of countries is almost identical for the NoMoney and the Money condition (Spearman's $\rho = 0.939$, $P < 0.001$), we pooled data between the two conditions. All regressions control for treatment condition, recipient and situational characteristics, as well as institution fixed effects. Fig. S5 presents the corresponding coefficients and standard errors (adjusted for clustering at the country-level). We standardized the explanatory variables to have a mean of zero and a standard deviation of one, so the coefficients can be interpreted as the difference in reporting rates associated with a one standard

[11] Some of the variables were not available for all countries in our dataset. Where possible, we updated the data to obtain better geographic coverage. For example, measures of historic institutions were substituted from predecessor countries and we manually coded linguistic traits for several countries using the *World Atlas of Language Structures (WALS)*. Fig. S7 shows that the results are qualitatively similar if we only use data from the original sources.

deviation change in the explanatory variable. To account for multiple hypothesis testing, we report P-values adjusted for the false discovery rate (*26*). Figs. S8 and S9 show that our results are robust when we conduct our regression analysis separately for the Money and NoMoney conditions.

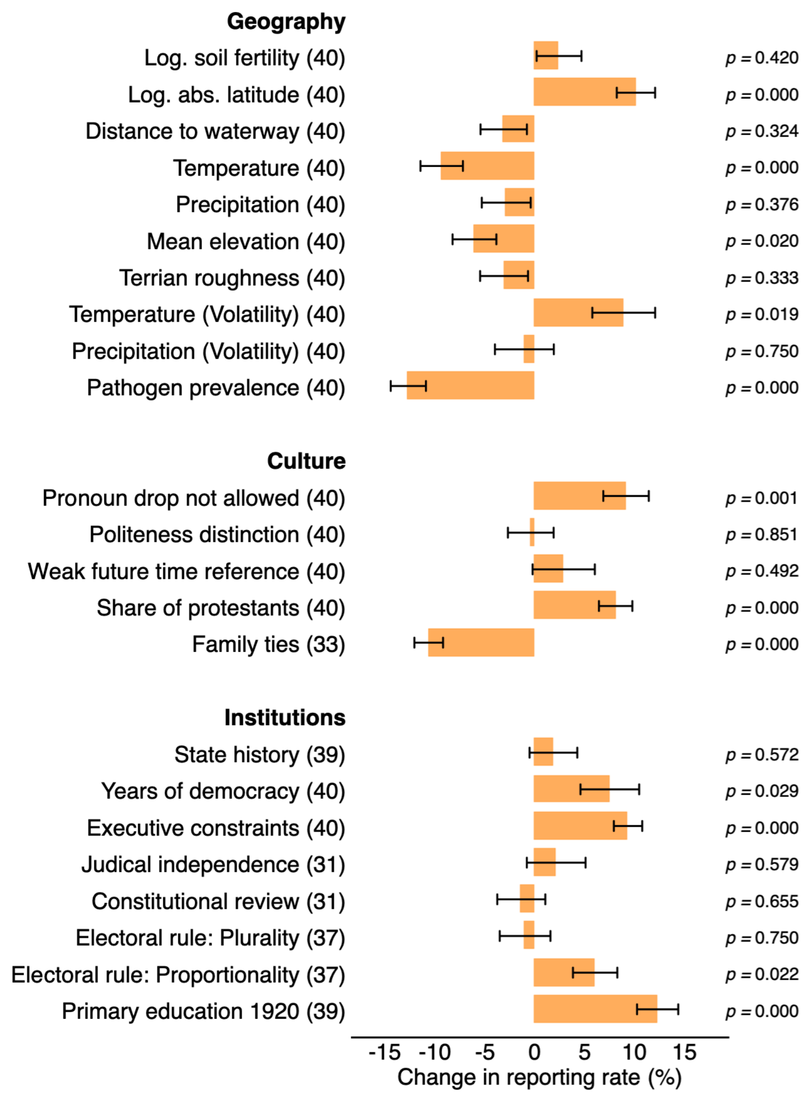
We first examined whether rates of civic honesty are correlated with commonly-discussed geographic conditions: soil fertility, absolute latitude, distance to waterways, temperature, precipitation, elevation, and terrain ruggedness. These geographic conditions have been found to foster economic development (*43, 61*), and we find that such variables are also significantly associated with civic honesty. Country-level reporting rates for lost wallets were associated with absolute latitude ($t_{39} = 5.26$, $P < 0.001$), lower temperature ($t_{39} = 4.40$, $P < 0.001$), and lower elevation ($t_{39} = 2.77$, $P = 0.020$). These findings suggest that civic honesty may be a channel through which geography affects economic development, in that geographic conditions and climate could have influenced the scope of social interactions and cooperation in pre-industrial societies. Norms of trust and cooperation may have in turn facilitated the transition from agricultural societies to market economies, which are based on interactions with out-group members and strangers (*62–68*). Another possibility is that geography indirectly influences civic honesty by promoting favorable economic conditions, which in turn increases rates of honesty (*69–71*).

We next examined the role of historical weather variability. Buggle and Durante (*72*) advanced the hypothesis that subsistence farmers developed persistent norms of cooperation and trust in strangers to cope with climate risk, which in turn facilitated exchanges between communities or helped to establish geographically-diversified insurance agreements (*73–75*). Using regional survey data from Europe, they found that historical weather variability is positively correlated with trust. Corroborating Buggle and Durante's survey results, we find that historical seasonal variability in temperature is also positively correlated with reporting rates in our study ($t_{39} = 2.82$, $P = 0.019$).[12]

We conclude our analysis of geographic factors by examining the relationship between historical prevalence of infectious diseases and civic honesty. According to the prominent pathogen-stress

[12] We also do not observe a significant correlation between precipitation and reporting rates ($t_{39} = 0.36$, $P = 0.750$). According to (*76*), temperature shocks were more decisive for productivity in the pre-industrial era than precipitation.

Fig. S5. Correlates of civic honesty



Notes: OLS coefficient estimates with standard errors clustered at the country-level. The dependent variable takes on the value 100 when an individual reported a wallet and 0 otherwise. Each coefficient has been estimated separately using standardized explanatory variables. They can therefore be interpreted as the difference in reporting rates associated with a one standard deviation change in the explanatory variable. We control for treatment status, institution fixed effects, and our standard set of control variables for recipient and situational characteristics: dummies for age above 40 years and gender, as well as the presence of a computer, coworkers, and other bystanders. To correct for multiple hypothesis testing, *P*-values are adjusted for false discovery rate (*26*). The number of countries included in the regressions is indicated in parentheses.

theory of sociality, communities that lived in regions with high exposure to infectious diseases were less likely to interact with strangers to prevent potential infection of novel pathogens, and as a result adopted collectivistic norms limited to one's immediate in-group (*77, 78*). Given that the lost wallets in our study always belonged to a stranger, recipients in locations with historically high pathogen prevalence may have felt less compunction to return a lost wallet to an out-group member. Consistent with this hypothesis, we find a sizable negative association between historical pathogen prevalence and civic honesty ($t_{39} = 7.20$, $P < 0.001$).

We next explored the relationship between civic honesty and cultural proxies for a generalized sense of morality — that is, moral norms and obligations that extend beyond one's in-group to anonymous strangers (*79*). To do so we first examined the role of different language structures, as language is thought to directly shape norms and expectations about behavior. For instance, Kashima and Kashima (*38*) proposed that languages which do not permit the dropping of first person pronouns (e.g., "I" in English or "ich" in German) serve to demarcate an individual from his or her social context, in turn reinforcing values around individual autonomy and responsibility. We found a strong positive correlation between reporting rates and countries with languages which do not permit the dropping first personal pronouns ($t_{39} = 4.00$, $P < 0.001$). This finding is consistent with prior work demonstrating that individualistic values are positively related to behaviors in line with generalized morality norms (*37, 80*). By contrast, we failed to find a reliable correlation between reporting rates and the use of multiple second person pronouns ($t_{39} = 0.19$, $P = 0.851$) or weak future time reference ($t_{39} = 0.92$, $P = 0.492$), two linguistic features that have received attention in the literature.[13]

Moving away from language to other cultural proxies of generalized morality, we next explored Protestantism. A long-standing literature in sociology and political science (*83, 84*) argues that Protestantism is conducive to social capital, and we find that countries with a higher share of Protestants also exhibit significantly more honest behavior ($t_{39} = 4.82$, $P < 0.001$). This is in

[13] Usage of multiple second person pronouns (e.g., "tu" and "vous" in French) as politeness distinction has been postulated to make status hierarchy and social distance more salient between speakers (*38, 81*). The weak future time reference feature allows the speaker to use the same grammatical tense to talk about future and present events and has been linked to greater patience and less impulsive behaviors (*39, 82*).

line with prior work finding that Protestantism encourages applying the same behavioral standards to in-group and out-group members, leading to higher trust in strangers (*22, 23, 85–87*). Indeed, we also found that stronger family ties are negatively correlated with reporting rates ($t_{32} = 7.42$, $P < 0.001$), as stronger family ties imply norms of cooperation that are often limited to one's narrow in-group (*41, 88–90*).

For the final part of our analysis, we explored some of the institutional determinants of civic honesty. The theoretical and empirical literature has examined both the complementarity between state formation and civic behavior (through the internalization of formal rules and increased trust in institutions), and their substitutability (as formal institutions may also crowd-out civic behavior) (*91–96*). We failed to find a significant association between state history — a commonly-used index of experience with formal government institutions (*42*) — and civic honesty ($t_{38} = 0.77$, $P = 0.572$). However, we found that both historical experience with democratic institutions and political constraints on executive power are positively correlated with reporting rates ($t_{39} = 2.55$, $P = 0.029$ for democratic history; $t_{39} = 6.54$, $P < 0.001$ for political constraints). This is consistent with the hypothesis that inclusive political institutions and the prevention of abuses of power are essential for civic behavior (*84, 88*).

Some researchers have argued, however, that commonly-used measures of societal institutions are potentially problematic because they measure time-varying political outcomes rather than permanent constraints (*44*). To address this concern we also analyzed a country's electoral rules (i.e., plurality and proportionality) and judicial checks and balances (i.e., judicial independence and constitutional review), which tend to be relatively time-invariant. Electoral systems based on plurality rule are thought to promote accountability due to the winner-take-all character of electoral competition,[14] but at the cost of targeting benefits to narrow constituencies and less overall representativeness (*97*). Proportional representation, on the other hand, is thought to be more inclusive and promotes broader democratic consensus.[15] Using data from Beck *et al.* (*46*), we found that countries

[14]The US and the UK are prime examples of countries with a plurality system where geographically defined constituencies elect one representative each.

[15]Examples of proportional representation include Scandinavian countries where each constituency elects several representatives. In these countries additional mechanisms are in place to ensure that the allocation of seats closely

with proportional representation exhibit significantly higher reporting rates ($t_{36} = 2.71$, $P = 0.022$), while plurality representation is not reliably related to civic honesty ($t_{36} = 0.40$, $P = 0.750$). These results suggest that broad political representation could be a key factor underlying the correlation between democratic institutions and civic honesty. We also used judicial independence and constitutional review as constitutional measures of the judiciary's power to constrain the executive. While these measures have been associated with political and economic freedom in previous studies (*86*), we failed to observe a significant correlation with reporting rates[16] ($t_{30} = 0.72$, $P = 0.579$ for judicial independence; $t_{30} = 0.58$, 0.655 for constitutional review).

Our last institutional variable involves national education. The history of national education is closely intertwined with the formation of the modern state (*99, 100*), so we examined the relationship between historical primary school enrollment rates and civic honesty. It has been argued that socialization is crucial to most primary education curricula and serves to ease interactions with strangers (*101*). We observed a significant and sizable positive correlation between historical rates of primary education and civic honesty ($t_{38} = 5.95$, $P < 0.001$), consistent with the hypothesis that education contributes to the formation of social capital (*2, 22, 79, 102–104*).

Given that geography has been linked to culture and institutions (*105–109*), it is possible that the correlations we observe between civic honesty and institutional variables may be spurious when not controlling for geographic conditions. We examined the robustness of our results to this concern by controlling for the first principal component of all geographic variables, and found qualitatively similar results[17] (see Fig. S10). The first principal component of our set of geographic variables accounts for roughly 32% of the variance in civic honesty, and the first principal compo-

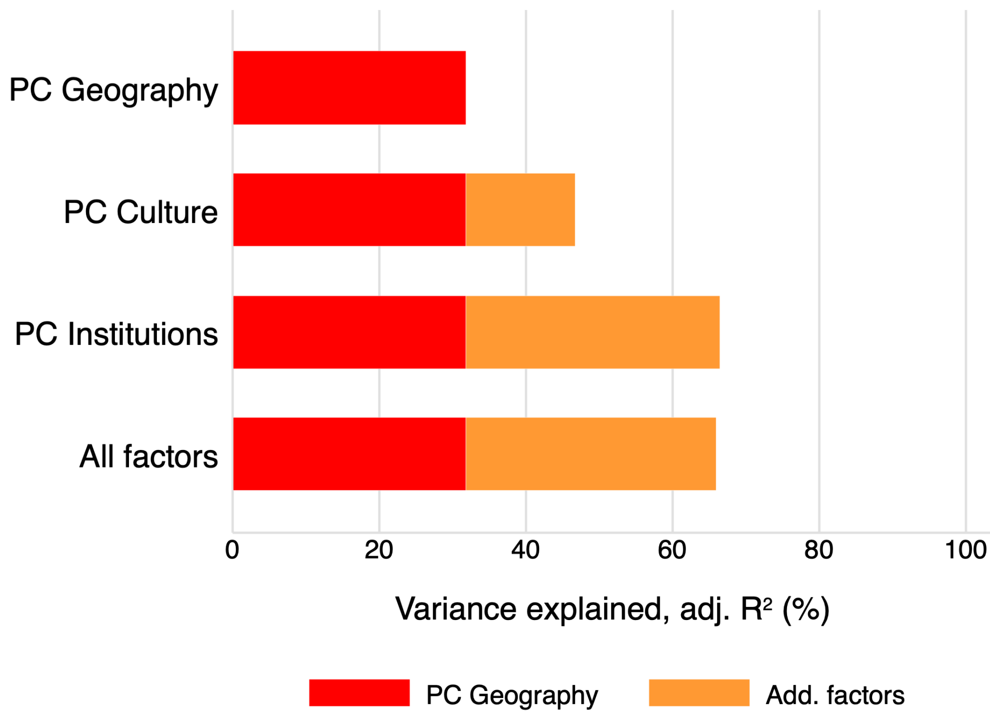
mirrors the overall popular vote. However, plurality and proportional representation are not mutually exclusive. Elements of both systems can coexist if a country's constitution stipulates different rules for electing representatives in a two-chamber legislature (e.g., Switzerland) or if proportional representation is combined with some sort of bonus for the winning party, as is the case in Italy (*98*).

[16]We note that for this analysis the sample is reduced to 31 countries due to data availability.

[17]Our results are similar if we control for the first three principal components or if the principal components are constructed using only the basic geographic factors, including soil fertility, absolute latitude, distance to waterway, temperature, precipitation elevation, and terrain ruggedness. As an alternative to controlling for the first principal component of geography, we also conducted the same regressions using the contemporary per capita income as our control variable. As shown in Fig S11, the results are largely unchanged.

nent of our set of cultural and institutional variables explains an additional 34% of the variation[18] (Fig. S6). Taken together, our analysis suggests that economically favorable geographic conditions, inclusive political institutions, national education, and cultural values that emphasize moral norms extending beyond one's in-group are positively associated with higher levels of civic honesty.

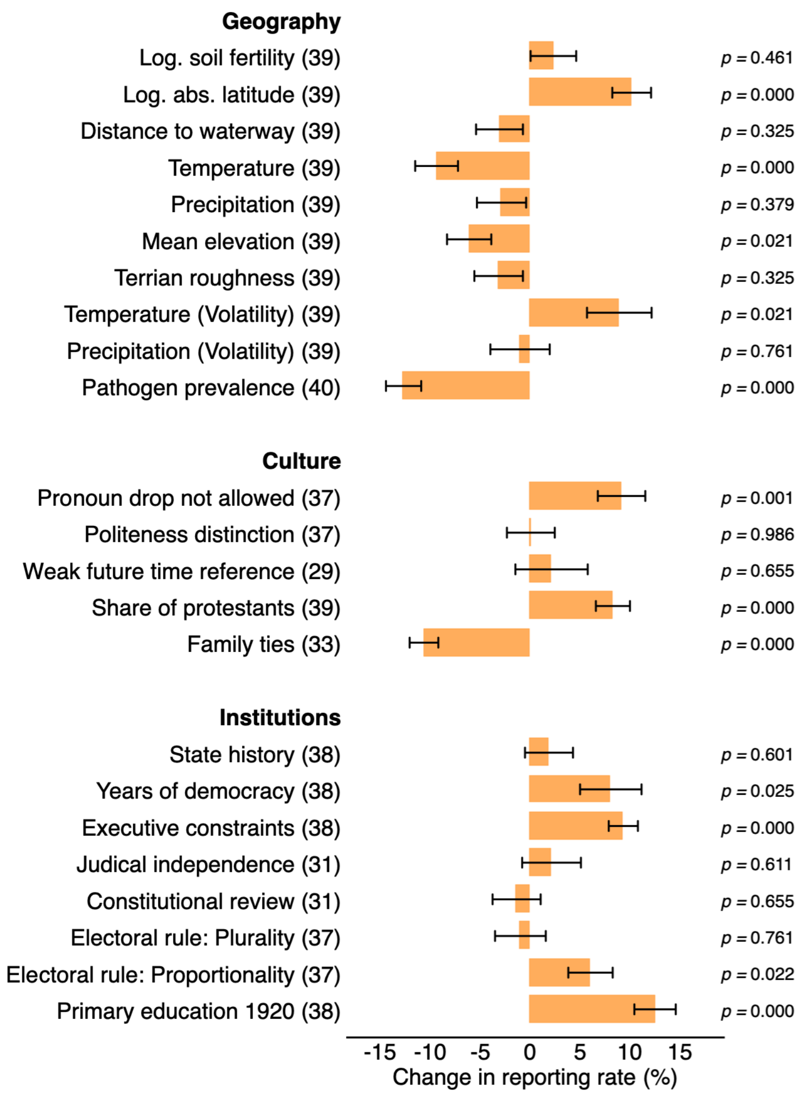
Fig. S6. Explaining cross-country variation



Notes: Explanatory power (adjusted R^2) of the first principal components of the geographic, cultural, and institutional variables. We regress country averages of regression-adjusted reporting rates (corrected for treatment indicators, institution fixed effects, and our standard set of control variables for individual characteristics and situational factors) on the first principal components of geography, geography and culture, geography and institutions, and all three categories together, respectively. To compute the first principal components of the variables in each category, we exclude variables with less than 37 observations (i.e., family ties, judicial independence, and constitutional review).

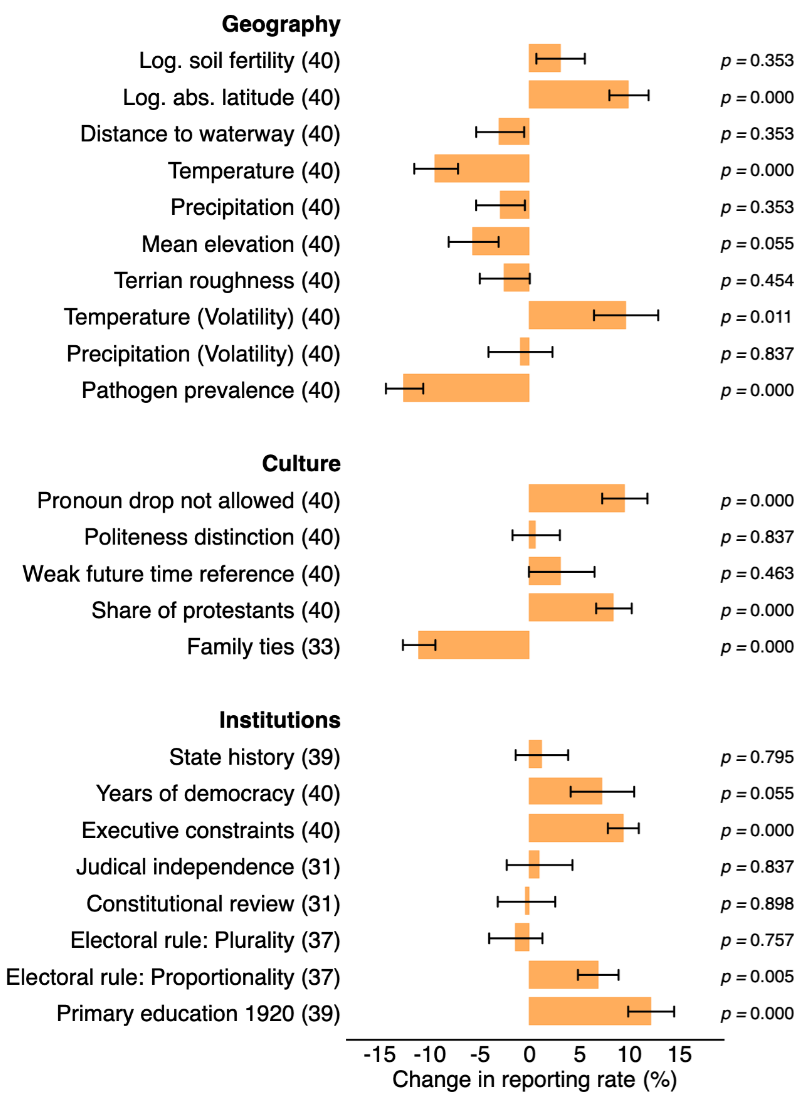
[18]To compute the first principal components for each category, we exclude variables with less than 37 observations (i.e., family ties, judicial independence, and constitutional review). The results are similar if we restrict the analysis to the 25 countries where all measures are available: Geography explains 40% of the variation in civic honesty and culture and institutions together explain an additional 25%.

Fig. S7. Correlates of civic honesty: original data only



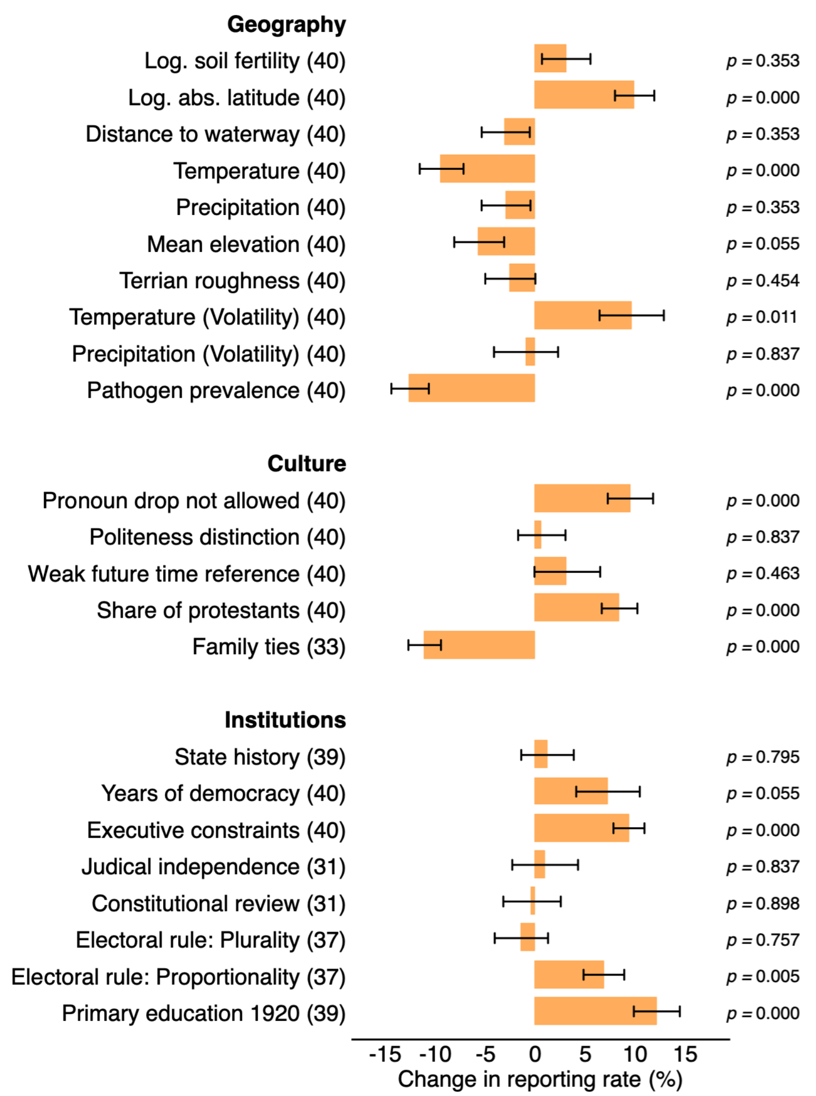
Notes: OLS coefficient estimates with standard errors clustered at the country-level. The dependent variable takes on the value 100 when an individual reported a wallet and 0 otherwise. Each coefficient has been estimated separately using standardized explanatory variables. They can therefore be interpreted as the difference in reporting rates associated with a one standard deviation change in the explanatory variable. We control for treatment status, institution fixed effects, and our standard set of control variables for recipient and situational characteristics: dummies for age above 40 years and gender, as well as the presence of a computer, coworkers, and other bystanders. To correct for multiple hypothesis testing, *P*-values are adjusted for false discovery rate (*26*). The number of countries included in the regressions is indicated in parentheses.

Fig. S8. Correlates of civic honesty: NoMoney



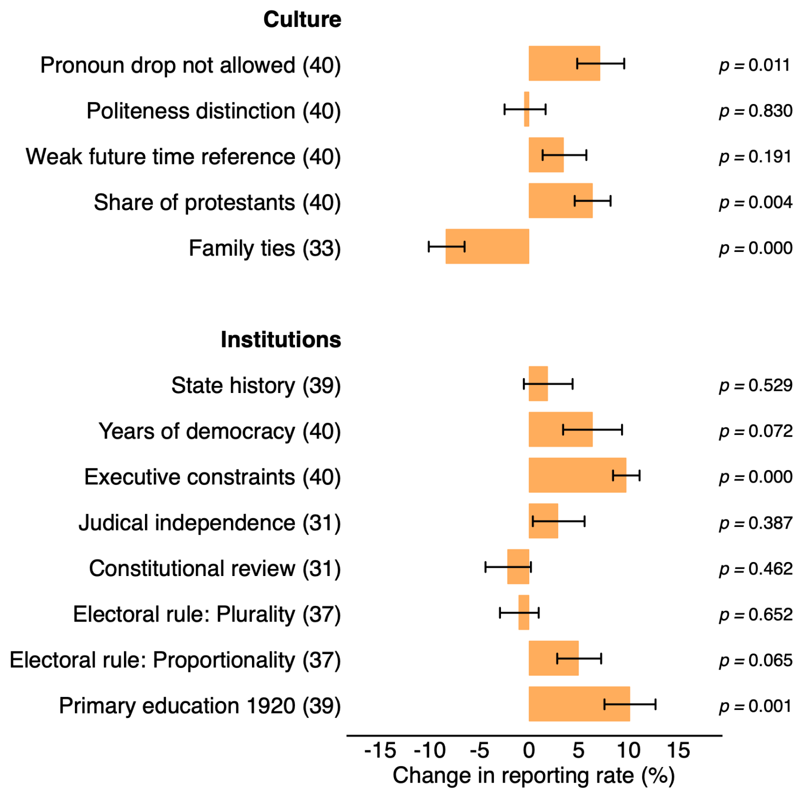
Notes: OLS coefficient estimates with standard errors clustered at the country-level. The sample is restricted to drop-offs in treatment NoMoney. The dependent variable takes on the value 100 when an individual reported a wallet and 0 otherwise. Each coefficient has been estimated separately using standardized explanatory variables. They can therefore be interpreted as the difference in reporting rates associated with a one standard deviation change in the explanatory variable. We control for treatment status, institution fixed effects, and our standard set of control variables for recipient and situational characteristics: dummies for age above 40 years and gender, as well as the presence of a computer, coworkers, and other bystanders. To correct for multiple hypothesis testing, *P*-values are adjusted for false discovery rate (*26*). The number of countries included in the regressions is indicated in parentheses.

Fig. S9. Correlates of civic honesty: Money



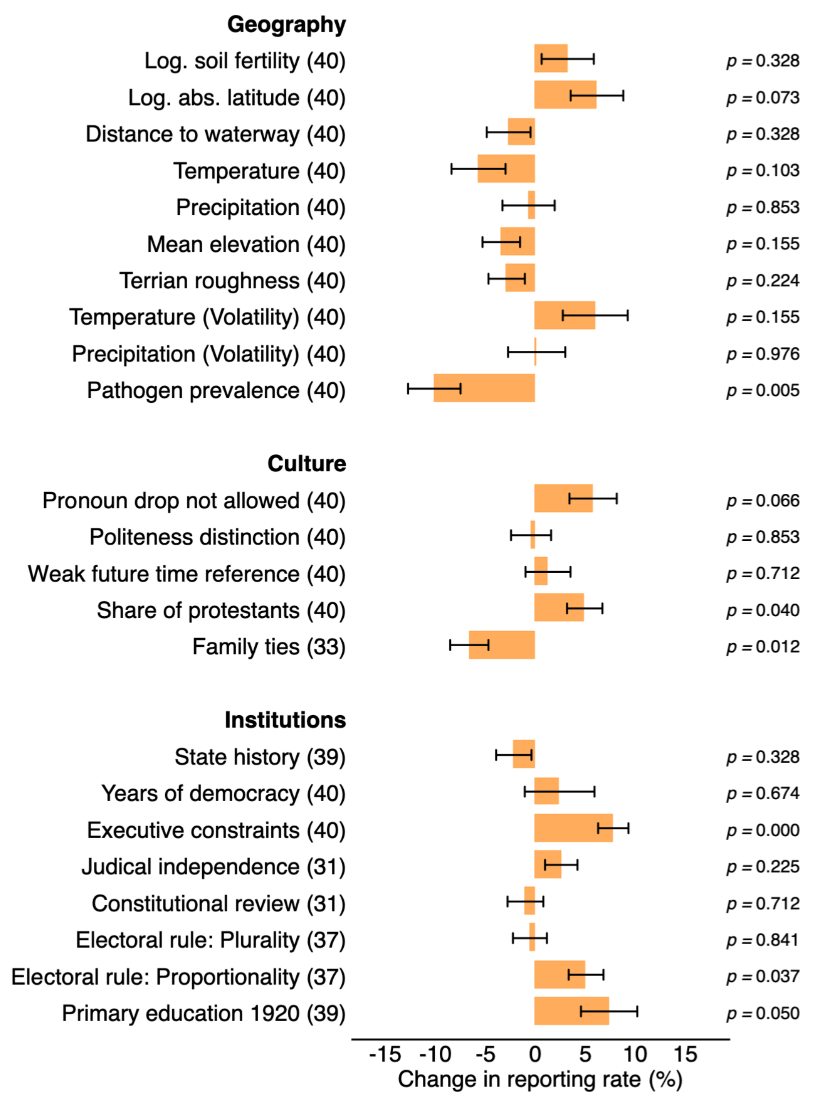
Notes: OLS coefficient estimates with standard errors clustered at the country-level. The sample is restricted to drop-offs in treatment Money. The dependent variable takes on the value 100 when an individual reported a wallet and 0 otherwise. Each coefficient has been estimated separately using standardized explanatory variables. They can therefore be interpreted as the difference in reporting rates associated with a one standard deviation change in the explanatory variable. We control for treatment status, institution fixed effects, and our standard set of control variables for recipient and situational characteristics: dummies for age above 40 years and gender, as well as the presence of a computer, coworkers, and other bystanders. To correct for multiple hypothesis testing, *P*-values are adjusted for false discovery rate (*26*). The number of countries included in the regressions is indicated in parentheses.

Fig. S10. Correlates of civic honesty: controlling for geography



Notes: OLS coefficient estimates with standard errors clustered at the country-level. The dependent variable takes on the value 100 when an individual reported a wallet and 0 otherwise. Each coefficient has been estimated separately using standardized explanatory variables. They can therefore be interpreted as the difference in reporting rates associated with a one standard deviation change in the explanatory variable. We control for the first principal component of all geographical measures, treatment status, institution fixed effects, and our standard set of control variables for recipient and situational characteristics: dummies for age above 40 years and gender, as well as the presence of a computer, coworkers, and other bystanders. To correct for multiple hypothesis testing, *P*-values are adjusted for false discovery rate (*26*). The number of countries included in the regressions is indicated in parentheses.

Fig. S11. Correlates of civic honesty: controlling for country GDP



Notes: OLS coefficient estimates with standard errors clustered at the country-level. The dependent variable takes on the value 100 when an individual reported a wallet and 0 otherwise. Each coefficient has been estimated separately using standardized explanatory variables. They can therefore be interpreted as the difference in reporting rates associated with a one standard deviation change in the explanatory variable. We control for the logarithm of a countries GDP per capita in 2010 (IMF World Economic Outlook; based on purchasing-power-parity), treatment status, institution fixed effects, and our standard set of control variables for recipient and situational characteristics: dummies for age above 40 years and gender, as well as the presence of a computer, coworkers, and other bystanders. To correct for multiple hypothesis testing, *P*-values are adjusted for false discovery rate (*26*). The number of countries included in the regressions is indicated in parentheses.

Supplementary Text 3: Alternative Explanations

We explored several alternative explanations for why rates of civic honesty tend to increase with greater amounts of money left in a wallet.

Fear of Punishment One possibility is that wallet recipients were concerned about possible punishment for not reporting the wallet, especially when a wallet contained relatively more money. We purposefully designed our experiment to minimize such concerns by telling recipients that the wallet was found on a different street and having our research assistants immediately leave upon handing over the wallet (thereby never receiving written confirmation for the lost item). We also note that lost property laws tend to be uncommon and even when in place are rarely enforced (*110*).[19]

We first address the issue of punishment concerns by exploiting regional variation in lost property laws within the US. The US legal system is based on common law, under which a person who finds lost property can keep the item until the original owner comes forward.[20] However, some states have enacted statutes that modify the common law's treatment of lost property. For instance, the state of New York imposes a fine of up to one hundred dollars if a finder willfully fails to report lost property.[21]

About half of our lost wallet observations in the US originate from states that have adopted statutes explicitly requiring finders to return lost property to the rightful owner or to a relevant agency, such as the police. We therefore divided our sample according to whether legal consequences could ensue for failing to return a lost wallet. If fear of legal punishment drives the in-

[19]In our representative survey we find a small but significant increase in self-reported fear of punishment with greater amounts of money in the wallet ($t_{2150} = 3.19$, $P = 0.001$, for the difference between the NoMoney and the Money condition; $t_{2150} = 2.45$, $P = 0.014$, for the difference between the Money and BigMoney condition). However, column 4 in Table S11 shows that while theft aversion concerns and altruism are positively correlated with the intention to report the wallet, self-reported fear of punishment does not predict the stated likelihood of reporting the wallet.

[20]Legal Information Institute, `https://www.law.cornell.edu/wex/lost_property`, accessed on September 18, 2016. Common law distinguishes between lost and mislaid property. Lost property is property that was unintentionally left behind by its owner. Mislaid property, on the other hand, is property that was intentionally set down in a location by its owner and then forgotten.

[21]See N.Y. Personal Property Law § 252 (3).

Tab. S14. Civic honesty and lost property laws

	Lost property law?	
	No	Yes
	(1)	(2)
Money	16.809**	20.350***
	(5.501)	(5.657)
BigMoney	29.887***	25.576***
	(5.876)	(6.245)
Constant	39.166**	35.011**
	(12.925)	(11.245)
Controls:		
Recipient	yes	yes
Situation	yes	yes
Institution FE	yes	yes
City FE	yes	yes
Other treatments	yes	yes
Money = BigMoney	0.026	0.404
Observations	496	504
Adjusted R^2	0.152	0.055

OLS estimates with robust standard errors in parentheses. Column 1 focuses on US states without a lost property law, whereas column 2 contains data from states with such a law. The dependent variable in both columns takes on the value 100 if a wallet was reported and 0 otherwise. "Money" and "BigMoney" are treatment indicators (we also include an indicator for treatment "Money-NoKey" but do not report its estimates for ease of exposition). Both models include binary control variables for individual and situational factors, including a recipient's age (above 40 years) and gender (male), as well as the presence of a computer, coworkers, and other bystanders. The models also include city and institution fixed effects. The bottom of the table reports P-values from t-tests for equality of the treatment coefficients. Significance levels: * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.

crease in reporting rates, then this relationship should be especially pronounced for states with lost property regulations. As shown in Table S14 however, we find similar treatment effects regardless of whether a state has a lost property law. Using seemingly unrelated regressions for states with and without property laws (*111*), we fail to find a reliable difference in the size of the coefficients between the two groups for either the Money treatment ($\chi^2_1 = 0.21$, $P = 0.646$) or the BigMoney treatment ($\chi^2_1 = 0.27$, $P = 0.607$). Thus, recipients in states with legal sanctions surrounding lost property did not act in a meaningfully different way from recipients in states without such laws.

A second way we address possible punishment concerns is by examining whether the presence of a security camera moderates our results. Security cameras could serve as proof that the wallet was turned in to the recipient and therefore amplify concerns about punishment if the wallet was not returned. After each drop-off, except in Poland and the United Kingdom, our research assistants took note of whether they observed a security camera. Column 1 in Table S15 shows that if

Tab. S15. Civic honesty and presence of security cameras

	Full sample	Security camera? No	Security camera? Yes
	(1)	(2)	(3)
Money	10.558***	8.963***	11.591***
	(0.732)	(1.167)	(0.950)
Security Camera	−2.659**		
	(0.956)		
Constant	40.096***	27.774*	38.699***
	(5.143)	(11.485)	(5.691)
Recipient	Yes	Yes	Yes
Situation	Yes	Yes	Yes
Institution FE	Yes	Yes	Yes
City FE	Yes	Yes	Yes
Other treatments	Yes	Yes	Yes
Observations	15369	5806	9563
Adjusted R^2	0.189	0.224	0.170

OLS estimates with robust standard errors in parentheses. Column 1 shows the estimates for the full sample as a benchmark, column 2 contains observations where no security camera was sighted, and column 3 includes only observations where a camera was sighted. The dependent variable in all models takes on the value 100 if a wallet was reported and 0 otherwise. "Money" is a dummy for treatment Money (we also include indicators for treatments "Money-NoKey" and "BigMoney" but do not report their estimates for ease of exposition). All models include binary control variables for recipient and situational characteristics, including a recipient's age (above 40 years), gender (male), and the presence of a computer, coworkers and other bystanders. The models also include city and institution fixed effects. Note that the sample does not include data from the United Kingdom and Poland because we did not collect data on security cameras. Significance levels: * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.

anything, the presence of a security camera during the drop-off lowered the likelihood of reporting a wallet by 2.7 percentage points ($t_{15044} = -2.78$, $P = 0.005$). While the treatment effect in the Money condition, relative to the NoMoney condition, is slightly larger for drop-off locations with cameras than those without ($\chi^2_1 = 3.20$, $P = 0.074$ when comparing the coefficient of Money in columns 2 and 3), the treatment effect is large and significant for both subsamples ($t_{5485} = 7.68$, $P < 0.001$ for column 2; $t_{9241} = 12.20$, $P < 0.001$ for column 3).

A third approach we use to address punishment concerns involves the presence of other individuals when performing a wallet drop-off. Recipients may have been worried about negative reactions from bystanders — an informal punishment — for not reporting a wallet. After performing the wallet drop-offs, our research assistants also noted whether coworkers and other individuals were present during the exchange. If worries about informal sanctions influenced recipient's behavior then we should observe smaller treatment effects when other individuals were not present.

Tab. S16. Civic honesty and social monitoring

	Full sample	No coworkers	No bystanders	Alone
	(1)	(2)	(3)	(4)
Money	10.792***	9.944***	10.083***	8.824***
	(0.712)	(0.884)	(1.216)	(1.506)
Constant	33.302**	28.147*	64.166**	67.158**
	(11.112)	(11.407)	(24.972)	(24.791)
Controls:				
Recipient	yes	yes	yes	yes
Situation	yes	yes	yes	yes
Institution FE	yes	yes	yes	yes
City FE	yes	yes	yes	yes
Other treatments	yes	yes	yes	yes
Observations	17295	11528	5939	4079
Adjusted R^2	0.185	0.178	0.205	0.196

Notes: OLS estimates with robust standard errors in parentheses. Column 1 shows the estimates for the full sample as a benchmark, column 2 includes observations without coworkers present, column 3 includes observations without other bystanders present, and column 4 includes observations where neither coworkers nor other bystanders were present. The dependent variable in all models takes on the value 100 if a wallet was reported and 0 otherwise. "Money" is a dummy for treatment Money (we also include an indicators for treatments "Money-NoKey" and "BigMoney" but do not report their estimates for ease of exposition). All models include binary control variables for recipient and situational characteristics, including a recipient's age (above 40 years), gender (male), and the presence of a computer. The models also include city and institution fixed effects. Significance levels: * $P < 0.05$, ** $P \leq 0.01$, *** $P < 0.001$.

Table S16 displays the results for the full sample compared to instances when no coworkers were present, no bystanders were present, and when the recipient and research assistant were completely alone during the exchange. Relative to the full sample, we fail to find a reliable difference in treatment effects when co-workers are not present ($\chi^2_1 = 0.56$, $P = 0.455$ when comparing the coefficient of Money in columns 1 and 2), when other individuals are not present ($\chi^2_1 = 0.25$, $P = 0.615$ comparing columns 1 and 3), and when recipients were alone ($\chi^2_1 = 1.40$, $P = 0.238$ comparing columns 1 and 4). We observe roughly similarly-sized treatment effects between the full sample and all subsamples, suggesting that the presence of others did not qualify our results.

Returning the Wallet but Pocketing the Money Another explanation for our main result is that recipients in the Money and BigMoney conditions may have been more likely to return the wallet after first pocketing the money. We decided not to collect reported wallets to minimize the inconvenience to the recipients. It is possible that some recipients contacted the owner to return the wallet without the money.

To examine this possibility we picked up all reported wallets in seven cities across the Czech Republic (82 wallets) and Switzerland (90 wallets). We selected these two countries because they differ markedly in their level of corruption and presumably also in dishonest behavior.[22] If some recipients reported the wallet after first pocketing the money, then we should observe wallets that are returned without any money (especially in the Czech Republic where corruption is more prevalent). However, we recovered 99% and 98% of the money from the wallets that we picked up in Switzerland and the Czech Republic, respectively, and we observe no reliable difference between the two countries ($Z = 0.22$, $P = 0.823$ by a rank-sum test). This suggests that collecting emails was a valid method to measure whether people would return a wallet with all of its contents.

Possible Finder's Fee for Returning a Wallet Another explanation for the increase in civic honesty for wallets with greater amounts of money is that the recipients expected a larger monetary reward (i.e., "finder's fee") when returning a wallet that contained relatively more money. To examine this possibility, we asked respondents in our representative survey experiments about their beliefs regarding a finder's fee and find results that are inconsistent with the behavioral patterns from our field experiments.

In the representative survey experiments, we asked the participants to estimate the likelihood that they would receive a financial reward from the owner, and if they received such a reward, how much money did they think they would get. We constructed a measure of *expected reward* by multiplying these two estimates together, and to facilitate comparability across countries we converted amounts to US dollars using the same exchange rate as in our field experiments. Overall, 42% of the participants stated that they would not expect a financial reward at all. The median expected reward ranged between US \$0.00 (Money-NoKey condition) and \$1.58 (High-Stakes condition) — cash amounts that were much lower than what the recipients could have gained from keeping the wallet (except for the NoMoney condition). Finally, we do not observe that the expected reward increased monotonically with the amount of money in the wallet, as shown in column 1 of Table S17. In fact, on average participants expected the highest reward in the

[22] In 2013, Transparency International ranked Switzerland 7th and the Czech Republic 57th out of 177 countries.

Tab. S17. Civic honesty and beliefs about finder's fees

	Expected reward (in US \$)	Reporting likelihood (in %)
	(1)	(2)
Money	−1.833***	2.417*
	(0.511)	(0.998)
BigMoney	−0.128	3.848***
	(0.460)	(0.976)
Money-NoKey	−2.816***	−2.428*
	(0.397)	(1.184)
Expected reward (in US \$)		0.009
		(0.039)
Constant	3.995***	86.378***
	(0.516)	(1.249)
Controls:		
Institution FE	yes	yes
Country FE	yes	yes
Money = BigMoney	0.000	0.125
Money = Money-NoKey	0.008	0.000
Observations	2160	2160
Adjusted R^2	0.028	0.029
F	13.320	6.619

Notes: OLS estimates with robust standard errors in parentheses. In column 1, the dependent variable is participants' expected financial reward for reporting the wallet (in US dollars). The dependent variable in column 2 is the likelihood that participants would report the wallet (as a percentage). "Money," "BigMoney," and "Money-NoKey," are treatment indicators. All models include country and institution fixed effects. The bottom of the table reports P-values from t-tests for equality of the treatment coefficients. Significance levels: * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.

NoMoney condition.[23] We also do not find that a higher expected reward is associated with a higher stated likelihood of reporting the wallet, as shown in column 2 of Table S17. Moreover, controlling for a respondent's expected reward does not meaningfully change our observed treatment effects (see column 2 of Table S11 for comparison). Overall, the prospect of a financial reward is unlikely to explain the monotonic increase in reporting rates.

[23]One potential explanation for this seemingly counterintuitive result is that the amount of cash in the Money condition serves as an upper bound on the amount people expect to receive as a finder's fee. Consistent with this interpretation, when examining conditional expectations about the reward (i.e., how much money a respondent expects to receive, conditional upon receiving a reward for returning the wallet) we find that only 12% of responses exceeded \$13.45 in the Money condition, compared to 35% of responses in the NoMoney condition ($z = 9.13$, $P < 0.001$). This difference is significant when examining each country (US, UK, and Poland) separately ($z > 4.00$ in all conditions, $P < 0.001$). Furthermore, and also consistent with a censoring effect, we observed greater variability in conditional expected finder fees in the NoMoney condition than in any other condition ($P < 0.001$ by a variance-ratio test for every pairwise comparison between the NoMoney condition and all other conditions).

Supplementary Text 4: Robustness Checks

Individual and Situational Factors To what extent do individual and situational factors drive cross-country differences in civic honesty? For instance, drop-off locations may have been more crowded in some countries with the possible consequence that recipients felt more observed and obliged to return the wallet. Or perhaps recipients were busier when there were more customers present during the drop-off and as a result less likely to report a wallet.

To examine the robustness of cross-country differences in civic honesty, we estimated the residuals from a regression that accounted for recipient and situational characteristics between locations as well as institution fixed effects. We conducted this analysis separately for the Money and NoMoney conditions, and then aggregated the residuals by country. For ease of exposition, we add the average reporting rate across all countries. The resulting regression-adjusted ranking and the original country ranking were virtually the same for both the NoMoney and Money conditions (Spearman's ρ = 0.976 and 0.990, respectively; both P-values < 0.001). Moreover, the range of reporting rates across countries remained large and almost identical when using the regression-adjusted data instead of the original data (Fig. S12). This suggests that differences in recipient and situational characteristics between locations did not account for large differences in civic honesty across countries.

Experimenter Effects We also examined the role of our research assistants in influencing recipient decisions to report the wallets. We used a total of 13 research assistants (all recruited from two German speaking universities), and purposely created overlaps for some of the countries they traveled to. We had two research assistants with overlapping presences in France, Germany, Italy, Malaysia, Poland, Spain, Switzerland, Turkey and the UK, and seven research assistants in the US. Table S18 presents an overview of the number of wallets each research assistant turned in by country. The numbers in parentheses represent the number of wallets for which there was at least one other research assistant performing drop-offs in the same city. These overlaps help us to distinguish between experimenter and city fixed effects.

We first explored the influence of research assistants by introducing experimenter fixed effects in our benchmark regression model. Tables S19 and S20 present the estimates of the treatment effects with and without experimenter fixed effects for each country where we had an overlap. We ran several tests to assess the influence of the research assistants. First, we found that the treatment effects in each country remained basically the same, regardless of whether we control for experimenter fixed effects.[24] Second, in the US we performed all 21 pairwise comparisons of the seven experimenter fixed effects and found that none of the comparisons are statistically significant at the 5% level (note that this is a conservative test since we do not adjust the P-values for multiple hypothesis testing). Third, we conducted joint significance tests of the experimenter fixed effects and found null results in all countries (F-tests in Tables S19 and S20). Finally, we computed the change in the variance explained (measured by the adjusted R^2) when we augment our benchmark specification with experimenter fixed effects and found virtually no change in the variance explained (as shown at the bottom of Tables S19 and S20). Overall, we find little evidence that differences between research assistants are driving our results.

Differences in Email Usage Since our measure of civic honesty relied on recipients contacting the owner by email, one concern is that differences in exposure to email communication could be responsible for cross-country differences in reporting rates. Yet, we focused on drop-off locations in urban places and included institutions where email communication is common. In particular, hotel staff should be able to communicate via email in all parts of the world. Consequently, if email experience is a key driver of differences in reporting rates, we should see substantially less heterogeneity when we restrict our sample to hotels. However, Fig. S13 shows that this is not the case. We still observe large differences in reporting rates across countries when focusing

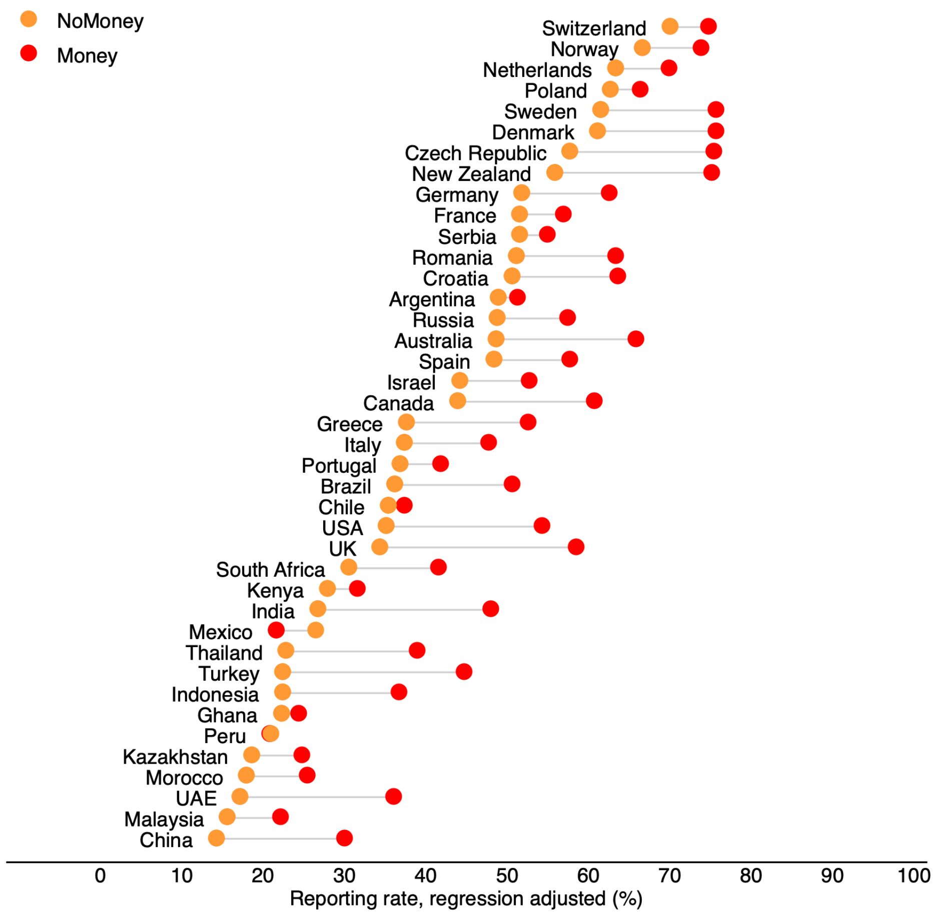
[24] We also estimated the same regression model as in column 2 of Table S8 and added the experimenters' age and gender as explanatory variables. Both coefficients failed to reach statistical significance, suggesting that experimenter age and gender did not reliably influence reporting rates among recipients ($t_{16924} = 1.17$, $P = 0.243$ for age; $t_{16924} = 1.18$, $P = 0.237$ for gender). We also failed to find a significant interaction effect between the gender of the experimenter and gender of the recipient ($t_{16923} = 0.90$, $P = 0.368$). However, these null results for experimenter gender should be interpreted carefully given that we only had two female research assistants.

on hotels only. As a further robustness check, we included the share of firms that use email to interact with their customers and suppliers in a country (from the World Bank Global Enterprise Survey) as an additional control variable to construct the regression-adjusted measure of civic honesty.[25] Fig. S14 shows that the differences between countries remain large, and the regression-adjusted ranking is almost identical to the unconditional ranking (Spearman's $\rho = 0.950$, $P < 0.001$ for the NoMoney condition, and $\rho = 0.932$, $P < 0.001$ for the Money condition). This suggests that experience with email communication was not a major driver of cross-country differences in reporting rates.

Differences in Economic Development We also assessed the extent that cross-country variation in civic honesty was robust when controlling for differences in economic development. For this purpose, we included contemporary per capita income as an additional control variable for the estimation of regression-adjusted reporting rates. The results in Fig S15 demonstrate that cross-country differences remain substantial, even when controlling for economic development. The regression-adjusted rankings from Fig S15 are also positively correlated with the unconditional rankings from Fig. 1 (Spearman's $\rho = 0.705$, $P < 0.001$ for the NoMoney condition; Spearman's $\rho = 0.753$, $P < 0.001$ for the Money condition).

[25]The Global Enterprise Survey does not cover most Western European countries and North America, so we limit our analysis of email usage to 27 countries.

Fig. S12. Regression-adjusted ranking



Notes: Regression-adjusted share of wallets reported in the NoMoney (US \$0) and Money (US \$13.45) condition by country. We regress individual decisions to report a wallet on recipient (age and gender of the recipient) and situational control variables (presence of a computer, number of coworkers and other bystanders) as well as institution fixed effects, and subsequently computed residuals for treatment Money and NoMoney. Finally, we aggregated residuals for each country and added the overall average reporting rate. The original and the regression-adjusted ranking are highly correlated for both the NoMoney and Money conditions (Spearman's $\rho = 0.976$, $P < 0.001$ and $\rho = 0.990$, $P < 0.001$, respectively).

Tab. S18. Drop-offs by experimenters and country: overlaps

Experimenter	Country									
	France	Germany	Italy	Malaysia	Poland	Spain	Switzerland	Turkey	United Kingdom	United States
#1		75	202				65	64		90
		(75)	(50)				(65)	(64)		(90)
#2		325				71				
		(76)				(53)				
#3	385						334			
	(111)						(63)			
#4						329			360	
						(57)			(107)	
#5										62
										(62)
#6										75
										(75)
#7				109						
				(109)						
#8				291						86
				(109)						(86)
#9	417		198		399				772	
	(111)		(46)		(42)				(96)	
#10					401					
					(42)					
#11								336		569
								(118)		(74)
#12										58
										(58)
#13										60
										(60)
Total obs.	802	400	400	400	800	400	399	400	1132	1000

Notes: Number of wallets turned in by experimenter and country in countries with overlaps. The number of drop-offs where at least one other experimenter turned in wallets in the same city is in parenthesis.

Tab. S19. Experimenter effects

	France		Germany		Italy		Malaysia		Poland	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Money	5.423	5.449	10.651*	10.629*	11.683*	11.567*	5.791	5.758	3.310	3.311
	(3.394)	(3.395)	(4.385)	(4.388)	(4.747)	(4.771)	(3.884)	(3.908)	(4.690)	(4.692)
BigMoney									11.761**	11.790**
									(4.410)	(4.422)
Constant	63.758***	63.393***	77.793***	82.424***	49.231***	45.81251**	41.396***	40.889***	59.380***	58.974***
	(11.491)	(11.508)	(9.841)	(10.131)	(14.620)	(17.275)	(10.701)	(11.401)	(11.216)	(11.198)
Controls:										
Experimenter FE	no	yes	no	yes	no	yes	no	yes	no	yes
Recipient	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Situation	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Institution FE	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
City FE	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Other treatment	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
F-test Experimenter effects		0.537		0.185		0.700		0.894		0.202
Money = BigMoney									0.058	0.058
Observations	802	802	400	400	400	400	400	400	794	794
Adjusted R^2	0.074	0.073	0.185	0.186	0.091	0.089	0.050	0.048	0.050	0.051

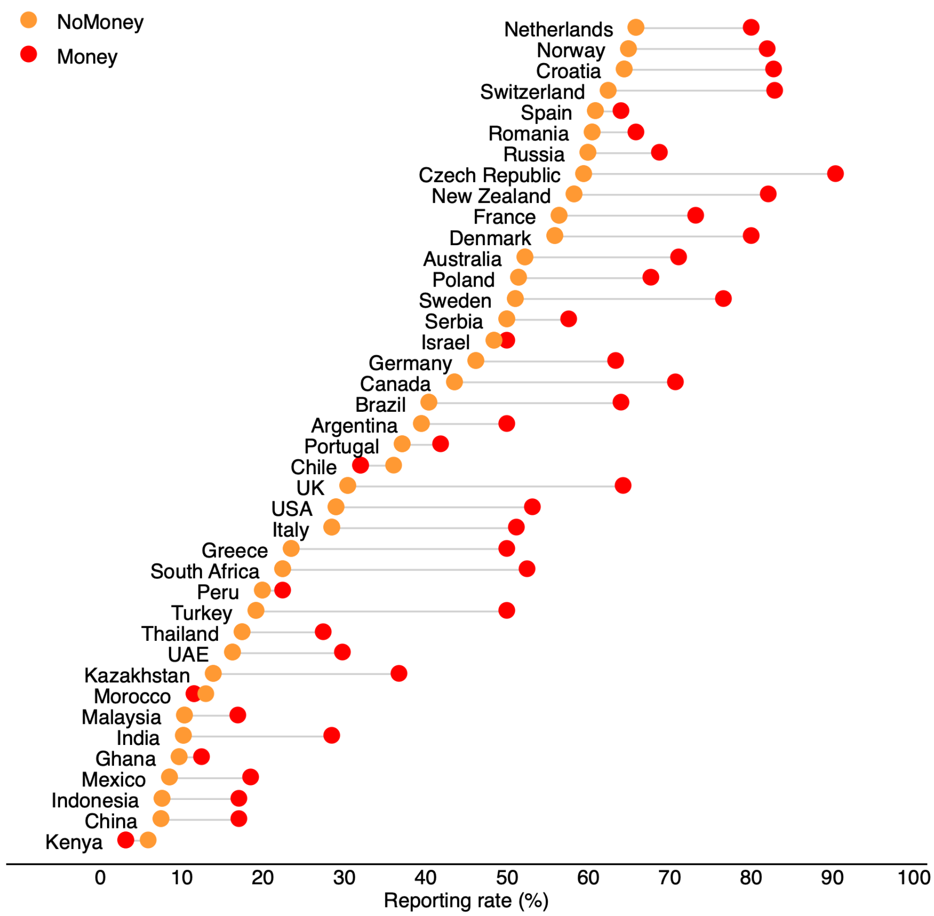
Notes: OLS estimates with robust standard errors in parentheses. The dependent variable in all models takes on the value 100 if a wallet was reported and 0 otherwise. "Money," "BigMoney," and "Money-NoKey" are treatment indicators. All models include binary control variables for individual and situational factors, including a recipient's age (above 40 years) and gender (male), the presence of a computer, other people, and coworkers. All models include city and institution fixed effects. Models in the even columns additionally control for experimenter fixed effects. The bottom of the table reports P-values from t-tests for equality of the coefficients of treatments "Money" and "BigMoney." Significance levels: * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.

Tab. S20. Experimenter effects (continued)

	Spain		Switzerland		Turkey		United Kingdom		United States	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Money	9.479	10.086*	4.094	4.194	20.087***	20.160***	23.106***	23.074***	18.301***	18.235***
	(4.899)	(4.920)	(4.178)	(4.187)	(4.486)	(4.492)	(3.851)	(3.851)	(3.934)	(3.955)
BigMoney							35.941***	35.929***	27.832***	27.793***
							(4.567)	(4.571)	(4.260)	(4.247)
Constant	30.454*	20.523	62.645***	66.160***	12.633	10.883	25.763**	28.239*	34.445**	39.464**
	(15.482)	(17.847)	(11.165)	(12.697)	(11.706)	(14.075)	(9.345)	(11.243)	(11.291)	(13.585)
Controls:										
Experimenter FE	no	yes	no	yes	no	yes	no	yes	no	yes
Recipient	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Situation	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Institution FE	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
City FE	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Other treatment	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
F-test Experimenter effects		0.271		0.579		0.819		0.699		0.587
Money = BigMoney							0.001	0.001	0.027	0.027
Observations	400	400	399	399	400	400	1132	1132	1000	1000
Adjusted R^2	0.046	0.047	0.056	0.054	0.086	0.083	0.122	0.122	0.100	0.099

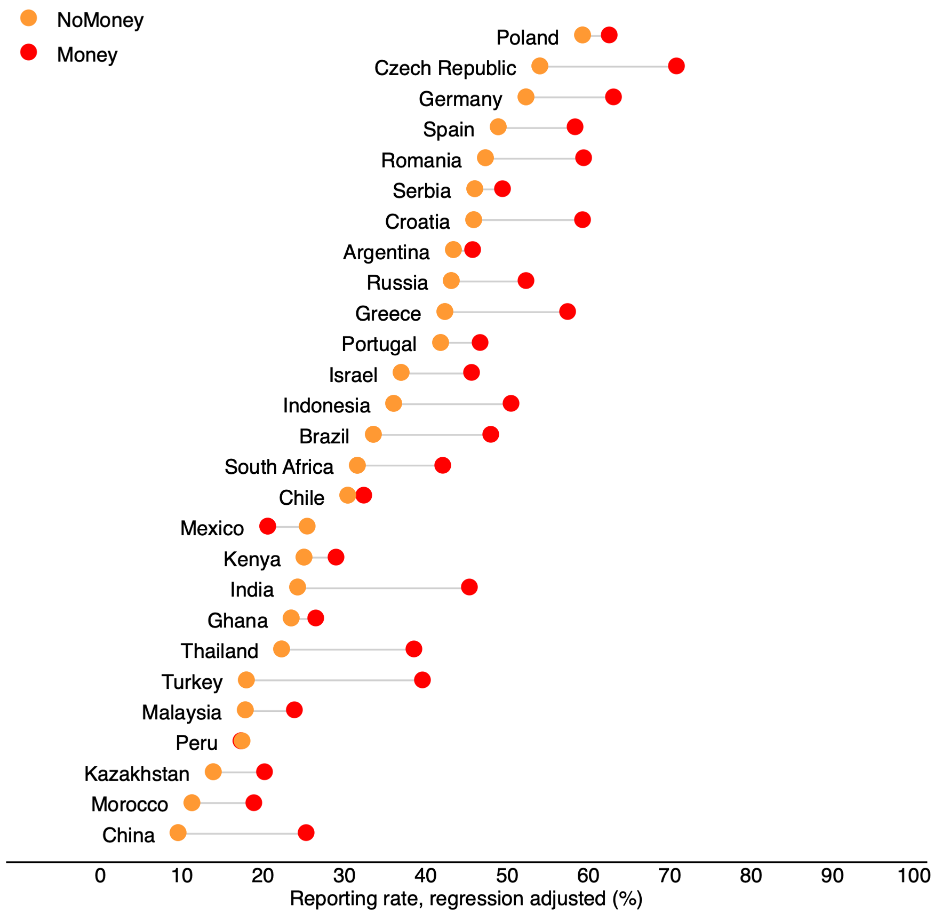
Notes: OLS estimates with robust standard errors in parentheses. The dependent variable in all models takes on the value 100 if a wallet was reported and 0 otherwise. "Money," "BigMoney," and "Money-NoKey" are treatment indicators. All models include binary control variables for individual and situational factors, including a recipient's age (above 40 years) and gender (male), the presence of a computer, other people, and coworkers. All models include city and institution fixed effects. Models in the even columns additionally control for experimenter fixed effects. The bottom of the table reports P-values from t-tests for equality of the coefficients of treatments "Money" and "BigMoney." Significance levels: * $P \leq 0.05$, ** $P < 0.01$, *** $P < 0.001$.

Fig. S13. Country ranking for hotels



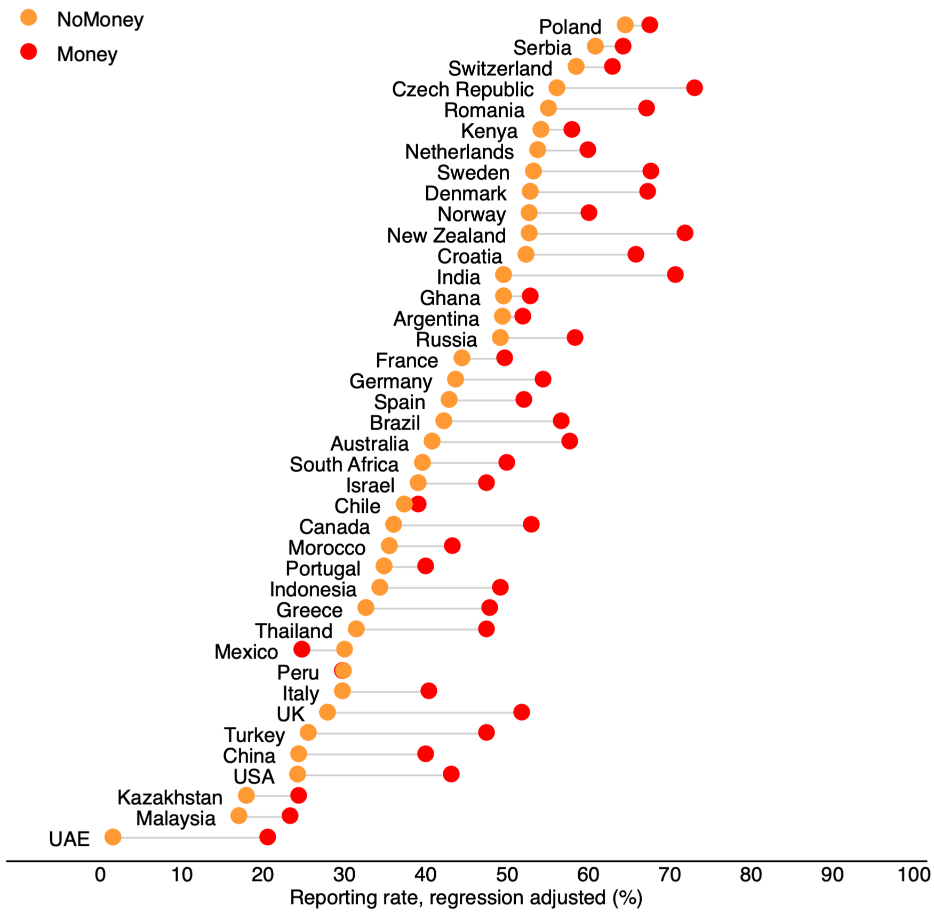
Notes: Share of wallets reported by hotel employees in treatments NoMoney (US $0) and Money (US $13.45) by country. The amount of money in the wallet is adjusted to purchasing power parity for each country. 'AVERAGE' shows the averages across all 40 countries.

Fig. S14. Regression-adjusted ranking: email usage



Notes: Regression-adjusted share of wallets reported in treatment decisions to report a wallet on the share of firms that use email to interact with their customers and suppliers in a country (from the World Bank Global Enterprise Survey), individual (age and gender of the recipient) and situational control variables (presence of a computer, number of coworkers and other bystanders) as well as institution fixed effects, and subsequently computed residuals for treatments Money and NoMoney. Finally, we aggregated residuals for each country and added the overall average reporting rate. The regression-adjusted ranking is almost identical to the unconditional ranking (Spearman's $\rho = 0.950$, $P < 0.001$ for treatment NoMoney, and $\rho = 0.932$, $P < 0.001$ for treatment Money). Due to missing data, the estimates are based on a sample of 27 countries.

Fig. S15. Regression-adjusted ranking: country GDP



Notes: Regression-adjusted share of wallets reported in the NoMoney (US \$0) and Money (US \$13.45) condition by country. We regress individual decisions to report a wallet on the logarithm of a countries GDP per capita in 2010 (IMF World Economic Outlook; based on purchasing-power-parity), in addition to recipient (age and gender of the recipient) and situational control variables (presence of a computer, number of coworkers and other bystanders) as well as institution fixed effects. We subsequently computed residuals for treatment Money and NoMoney. Finally, we aggregated residuals for each country and added the overall average reporting rate. The original and the GDP-adjusted ranking are significantly correlated for both the NoMoney and Money conditions (Spearman's $\rho = 0.705$, $P < 0.001$ and $\rho = 0.753$, $P < 0.001$, respectively).

References

26. Y. Benjamini, Y. Hochberg, Controlling the false discovery rate: A practical and powerful approach to multiple testing, *Journal of the Royal Statistical Society. Series B* **57**, 289–300 (1995).

27. I. M. Fund, *World Economic Outlook, April 2015* (International Monetary Fund Washington, DC, 2015).

28. Q. Ashraf, O. Galor, The 'Out of Africa' hypothesis, human genetic diversity, and comparative economic development', *American Economic Review* **103**, 1–46 (2013).

29. N. Ramankutty, J. A. Foley, J. Norman, K. McSweeney, The global distribution of cultivable lands: current patterns and sensitivity to possible climate change, *Global Ecology and Biogeography* **11**, 377–392 (2002).

30. S. Michalopoulos, The origins of ethnolinguistic diversity, *American Economic Review* **102**, 1508–39 (2012).

31. J. L. Gallup, J. D. Sachs, A. D. Mellinger, Geography and economic development, *International Regional Science Review* **22**, 179–232 (1999).

32. W. D. Nordhaus, Geography and macroeconomics: New data and new findings, *Proceedings of the National Academy of Sciences* **103**, 3510–3517 (2006).

33. R. Durante, Risk, cooperation and the economic origins of social trust: an empirical investigation (2010).

34. O. Galor, Ö. Özak, The agricultural origins of time preference, *American Economic Review* **106**, 3064–3103 (2016).

35. D. R. Murray, M. Schaller, Historical prevalence of infectious diseases within 230 geopolitical regions: A tool for investigating origins of culture, *Journal of Cross-Cultural Psychology* **41**, 99–108 (2010).

36. Y. Gorodnichenko, G. Roland, Culture, institutions, and the wealth of nations, *Review of Economics and Statistics* **99**, 402–416 (2017).

37. G. Tabellini, Institutions and culture, *Journal of the European Economic Association* **6**, 255–294 (2008).

38. E. S. Kashima, Y. Kashima, Culture and language: The case of cultural dimensions and personal pronoun use, *Journal of Cross-Cultural Psychology* **29**, 461–486 (1998).

39. M. K. Chen, The effect of language on economic behavior: Evidence from savings rates, health behaviors, and retirement assets, *American Economic Review* **103**, 690–731 (2013).

40. R. La Porta, F. Lopez-de Silanes, A. Shleifer, R. Vishny, The quality of government, *Journal of Law, Economics, and Organization* **15**, 222–279 (1999).

41. A. Alesina, P. Giuliano, The power of the family, *Journal of Economic Growth* **15**, 93–125 (2010).

42. V. Bockstette, A. Chanda, L. Putterman, States and markets: The advantage of an early start, *Journal of Economic Growth* **7**, 347–369 (2002).

43. E. Spolaore, R. Wacziarg, How deep are the roots of economic development?, *Journal of Economic Literature* **51**, 325–69 (2013).

44. E. L. Glaeser, R. La Porta, F. Lopez-de Silanes, A. Shleifer, Do institutions cause growth?, *Journal of Economic Growth* **9**, 271–303 (2004).

45. R. La Porta, F. Lopez-de Silanes, C. Pop-Eleches, A. Shleifer, Judicial checks and balances, *Journal of Political Economy* **112**, 445–470 (2004).

46. T. Beck, G. Clarke, A. Groff, P. Keefer, P. Walsh, New tools in comparative political economy: The database of political institutions, *World Bank Economic Review* **15**, 165–176 (2001).

47. A. Benavot, P. Riddle, The expansion of primary education, 1870-1940: Trends and issues, *Sociology of Education* **61**, 191–210 (1988).

48. M. H. Davis, Measuring Individual Differences in Empathy: Evidence for a Multidimensional Approach, *Journal of Personality and Social Psychology* **44**, 113–126 (1983).

49. D. L. Paulhus, Two-component models of socially desirable responding, *Journal of personality and social psychology* **46**, 598 (1984).

50. C. Engel, Dictator games: a meta study, *Experimental Economics* **14**, 583–610 (2011).

51. J. Andreoni, J. Miller, Giving according to GARP: An experimental test of the consistency of preferences for altruism, *Econometrica* **70**, 737–753 (2002).

52. R. Bénabou, J. Tirole, Incentives and prosocial behavior, *American Economic Review* **96**, 1652–1678 (2006).

53. G. Charness, M. Dufwenberg, Promises and partnership, *Econometrica* **74**, 1579–1601 (2006).

54. P. Battigalli, M. Dufwenberg, Guilt in games, *American Economic Review* **97**, 170–176 (2007).

55. E. L. Krupka, R. A. Weber, Identifying social norms using coordination games: Why does dictator game sharing vary?, *Journal of the European Economic Association* **11**, 495–524 (2013).

56. J. Andreoni, B. D. Bernheim, Social image and the 50–50 norm: A theoretical and experimental analysis of audience effects, *Econometrica* **77**, 1607–1636 (2009).

57. P. Shrout, N. Bolger, Mediation in experimental and nonexperimental studies: New procedures and recommendations, *Psychological Methods* **7**, 422–445 (2002).

58. A. Greif, *Institutions and the path to the modern economy: Lessons from medieval trade* (Cambridge University Press, 2006).

59. A. Bisin, T. Verdier, The economics of cultural transmission and socialization, *Handbook of Social Economics* (Elsevier, 2011), vol. 1, pp. 339–416.

60. B. Enke, Kinship, cooperation, and the evolution of moral systems, *Quarterly Journal of Economics* (forthcoming).

61. Q. Ashraf, O. Galor, Dynamics and stagnation in the Malthusian epoch, *American Economic Review* **101**, 2003–41 (2011).

62. E. Ostrom, *Governing the Commons* (Cambridge University Press, 1990).

63. A. Greif, Cultural beliefs and the organization of society: A historical and theoretical reflection on collectivist and individualist societies, *Journal of Political Economy* **102**, 912–950 (1994).

64. M. Woolcock, Social capital and economic development: Toward a theoretical synthesis and policy framework, *Theory and Society* **27**, 151–208 (1998).

65. J. Henrich, *et al.*, In search of Homo Economicus: behavioral experiments in 15 small-scale societies, *American Economic Review* **91**, 73–78 (2001).

66. J. Henrich, *et al.*, Markets, religion, community size, and the evolution of fairness and punishment, *Science* **327**, 1480–1484 (2010).

67. J. Mokyr, The institutional origins of the industrial revolution, *Institutions and Economic Performance*, E. Helpman, ed. (Harvard University Press, 2008), pp. 64–119.

68. A. Litina, Natural land productivity, cooperation and comparative development, *Journal of Economic Growth* **21**, 351–408 (2016).

69. E. Sharma, N. Mazar, A. L. Alter, D. Ariely, Financial deprivation selectively shifts moral standards and compromises moral decisions, *Organizational Behavior and Human Decision Processes* **123**, 90–100 (2014).

70. R. Fisman, P. Jakiela, S. Kariv, How did distributional preferences change during the great recession?, *Journal of Public Economics* **128**, 84–95 (2015).

71. M. Ananyev, S. Guriev, Effect of income on trust: Evidence from the 2009 crisis in Russia, *The Economic Journal* (forthcomming).

72. J. C. Buggle, R. Durante, Climate Risk, Cooperation, and the Co-Evolution of Culture and Institutions, *CEPR Discussion Paper* **12380** (2017).

73. J. S. Dean, *et al.*, Human behavior, demography, and paleoenvironment on the Colorado Plateaus, *American Antiquity* **50**, 537–554 (1985).

74. E. A. Cashdan, Coping with risk: Reciprocity among the Basarwa of Northern Botswana, *Man* **20**, 454–474 (1985).

75. D. Nettle, Explaining global patterns of language diversity, *Journal of anthropological archaeology* **17**, 354–374 (1998).

76. O. Galor, V. Savitskiy, Climatic roots of loss aversion, *CESifo Working Paper* (2018).

77. C. L. Fincher, R. Thornhill, D. R. Murray, M. Schaller, Pathogen prevalence predicts human cross-cultural variability in individualism/collectivism, *Proceedings of the Royal Society of London B: Biological Sciences* **275**, 1279–1285 (2008).

78. C. L. Fincher, R. Thornhill, Parasite-stress promotes in-group assortative sociality: The cases of strong family ties and heightened religiosity, *Behavioral and Brain Sciences* **35**, 61–79 (2012).

79. G. Tabellini, Culture and institutions: economic development in the regions of Europe, *Journal of the European Economic Association* **8**, 677–716 (2010).

80. A. N. Licht, C. Goldschmidt, S. H. Schwartz, Culture rules: The foundations of the rule of law and other norms of governance, *Journal of Comparative Economics* **35**, 659–688 (2007).

81. R. Brown, A. Gilman, The pronouns of power and solidarity, *Style in Language*, T. A. Sebeok, ed. (MIT Press, 1960), pp. 253–276.

82. A. Falk, *et al.*, Global evidence on economic preferences, *The Quarterly Journal of Economics* **133**, 1645–1692 (2018).

83. M. Weber, *The Protestant Ethic and the Spirit of Capitalism* (Allen and Unwin, 1930).

84. R. D. Putnam, R. Leonardi, R. Y. Nanetti, *Making Democracy Work: Civic Traditions in Modern Italy* (Princeton University Press, 1993).

85. B. Arruñada, Protestants and Catholics: Similar work ethic, different social ethic, *Economic Journal* **120**, 890–918 (2010).

86. R. La Porta, F. Lopez-de Silanes, A. Shleifer, R. W. Vishny, Trust in large organizations, *American Economic Review* **87**, 333 (1997).

87. E. M. Uslaner, *The Moral Foundations of Trust* (Cambridge University Press, 2002).

88. E. C. Banfield, *The Moral Basis of a Backward Society.* (Free Press, 1958).

89. J. S. Coleman, *Foundations of Social Capital Theory* (Cambridge, MA: Harvard University Press, 1990).

90. A. Alesina, P. Giuliano, Family ties, *Handbook of Economic Growth* (Elsevier, 2014), vol. 2, pp. 177–215.

91. P. Aghion, Y. Algan, P. Cahuc, A. Shleifer, Regulation and distrust, *Quarterly Journal of Economics* **125**, 1015–1049 (2010).

92. R. Bénabou, J. Tirole, Identity, morals, and taboos: beliefs as assets, *Quarterly Journal of Economics* **126**, 805-855 (2011).

93. G. Tabellini, The scope of cooperation: Values and incentives, *Quarterly Journal of Economics* **123**, 905–950 (2008).

94. A. Cassar, G. d'Adda, P. Grosjean, Institutional quality, culture, and norms of cooperation: Evidence from behavioral field experiments, *Journal of Law and Economics* **57**, 821–863 (2014).

95. L. Guiso, P. Sapienza, L. Zingales, Long-term persistence, *Journal of the European Economic Association* **14**, 1401–1436 (2016).

96. S. Lowes, N. Nunn, J. A. Robinson, J. L. Weigel, The evolution of culture and institutions: Evidence from the Kuba kingdom, *Econometrica* **85**, 1065–1091 (2017).

97. T. Persson, G. Tabellini, Constitutions and economic policy, *Journal of Economic Perspectives* **18**, 75–98 (2004).

98. T. Nannicini, A. Stella, G. Tabellini, U. Troiano, Social capital and political accountability, *American Economic Journal: Economic Policy* **5**, 222–50 (2013).

99. A. Green, Education and state formation., *Education and State Formation* (Palgrave Macmillan, London, 1990), pp. 76–110.

100. E. M. Uslaner, B. Rothstein, The historical roots of corruption: state building, economic inequality, and mass education, *Comparative Politics* **48**, 227–248 (2016).

101. E. L. Glaeser, G. A. Ponzetto, A. Shleifer, Why does democracy need education?, *Journal of Economic Growth* **12**, 77–99 (2007).

102. S. Knack, P. Keefer, Does social capital have an economic payoff? A cross-country investigation, *Quarterly Journal of Economics* **112**, 1251–1288 (1997).

103. K. Milligan, E. Moretti, P. Oreopoulos, Does education improve citizenship? Evidence from the United States and the United Kingdom, *Journal of Public Economics* **88**, 1667–1695 (2004).

104. J. F. Helliwell, R. D. Putnam, Education and social capital, *Eastern Economic Journal* **33**, 1–19 (2007).

105. S. L. Engerman, K. L. Sokoloff, Factor endowments, institutions, and differential paths of growth among new world economies, *How Latin America Fell Behind: Essays on the Economic Histories of Brazil and Mexico, 1800-1914*, S. Haber, ed. (Stanford, CA: Stanford University Press, 1997), pp. 260–304.

106. D. Acemoglu, S. Johnson, J. A. Robinson, Reversal of fortune: Geography and institutions in the making of the modern world income distribution, *Quarterly Journal of Economics* **117**, 1231–1294 (2002).

107. N. Nunn, L. Wantchekon, The slave trade and the origins of mistrust in Africa, *American Economic Review* **101**, 3221–52 (2011).

108. Q. Ashraf, O. Galor, Cultural diversity, geographical isolation, and the origin of the wealth of nations, *National Bureau of Economic Research* **17640** (2011).

109. A. Alesina, P. Giuliano, N. Nunn, On the origins of gender roles: Women and the plough, *Quarterly Journal of Economics* **128**, 469–530 (2013).

110. M. D. West, Losers: recovering lost property in Japan and the United States, *Law & Society Review* **37**, 369–424 (2003).

111. A. Zellner, An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias, *Journal of the American statistical Association* **57**, 348–368 (1962).